THE UNIVERSITY *of* EDINBURGH

Linking mitochondrial metabolism and macrovesicular steatosis in non-alcoholic fatty liver disease

Matthew Christopher Sinton

PhD – The University of Edinburgh – 2020

Contents

Declaration

I, the undersigned, confirm that the work submitted is my own, except where work which has formed part of joint-authored publications has been included. The contribution of the other authors to this work has been explicitly indicated below. I confirm that the appropriate credit has been given within the thesis where reference has been made to the work of others.

Matthew Christopher Sinton Date: 24th July 2020

Abstract

Non-alcoholic fatty liver disease (NAFLD) is the most prevalent form of liver disease globally, affecting approximately 88% of obese individuals. It is characterised by hepatic lipid accumulation >5%, which can lead to inflammation and the development of non-alcoholic steatohepatitis (NASH), but this process is not well understood. NASH can, in turn, progress to cirrhosis and hepatocellular carcinoma. Observational studies in rodents and humans suggest that NAFLD is associated with dysfunctional mitochondrial metabolism, namely increased levels of TCA cycle activity. However, it is unclear if/how this is linked to lipid accumulation in hepatocytes.

To overcome the limitations of studying mitochondrial metabolism in currently used models, I utilised a human pluripotent stem cell (hPSC) model of NAFLD. These hPSCs were differentiated to hepatocyte-like cells (HLCs) and challenged with a cocktail of high energy substrates: lactate, pyruvate and octanoic acid (LPO). To determine the extent to which this model is representative of human NAFLD, I characterised patterns of lipid accumulation, and showed the presence of macrovesicular steatosis following challenge with LPO, which was associated with increased expression of lipid droplet associated PLIN2. In humans presenting with NAFLD ~90% develop macrovesicular steatosis, demonstrating the utility of this system in modelling NAFLD. Further, I performed a transcriptomic analysis and identified dysregulation of a number of NAFLD-associated genes, including CYP7A1, HMGCS2, SLC2A4, PLIN2, PPARA, PPARG, PPARGC1A and FASN. Pathway enrichment analyses revealed dysregulation of genes associated with the TCA cycle and oxidative phosphorylation.

Analysis of respiration in steatotic HLCs revealed decreased maximal respiratory capacity in steatotic HLCs, indicating dysfunction in the electron transport chain (ETC), between respiratory complexes I-III. Stable isotopic tracing of TCA cycle-associated metabolites revealed a truncation of the cycle, with inhibited conversion of succinate to fumarate, confirming dysfunction of succinate dehydrogenase (respiratory complex II in the ETC). Mitochondria were able to bypass this truncation

by increasing purine nucleotide cycle (PNC) activity, leading to replenishment of fumarate. Experimental depletion of fumarate in steatotic cells reversed the macrovesicular steatosis observed in LPO-challenged HLCs, with addition of exogenous fumarate leading to partial reversal of the phenotype. This demonstrates, for the first time, that fumarate is a driver of macrovesicular steatosis.

Fumarate is also a negative regulator of dioxygenases, which include the ten-eleven translocation (TET) enzymes. The TETs oxidise 5-methylcytosine (5hmC) to 5-hydroxymethylcytosine (5hmC). To determine whether TET activity was impacted by fumarate accumulation, I performed DNA immunoprecipitation sequencing to measure 5hmC. This revealed global changes in 5hmC in steatotic HLCs but only showed a limited of number of changes in protein coding and promoter regions. Changes in promoter regions displayed a moderate negative correlation with gene transcription. To determine whether fumarate impacted on gene transcription more broadly, further RNA sequencing analysis was performed, following inhibition of the purine nucleotide cycle. This showed decreased expression of lipid droplet associated PLIN2, and pathway enrichment analyses revealed disruption of pathways associated with cellular response to oxidative stress and iron ion homeostasis.

Taken together, these studies provide evidence that in NAFLD, exposure to high energy substrates results in TCA cycle truncation, leading to compensatory rewiring of mitochondrial metabolism through increased purine nucleotide cycle activity. This compensatory rewiring, however, leads to the development of macrovesicular steatosis. Thus, using this model, I show for the first time that fumarate is a driver of lipid accumulation in NAFLD.

Lay Abstract

Non-alcoholic fatty liver disease (NAFLD) is the most common form of liver disease around the world, affecting almost 9 out of 10 people who are obese. The disease begins when the liver accumulates too much fat and stores this in structures called lipid droplets. Although high levels of fat in a cell can be toxic, storing fats in lipid droplets can protect the cell. However, having too much fat in the liver can lead to the development of inflammation. People with this inflammation are more likely to develop cirrhosis, where the liver becomes excessively scarred, and have an increased risk of developing liver cancer or needing a transplant. Currently, NAFLD can only be reversed through diet and exercise, or by having weight-loss surgery.

To study NAFLD, I used a cell-based model by using human liver cells treated with a combination of chemicals to make them accumulate fat. DNA is found in all cells and is made up of small subunits called genes. These genes act as instructions for making proteins, which are the building blocks of cells. The genes can be turned on or off but, like the volume dial on a stereo, they can also be turned up or down, telling a cell to make more or less of a protein. I found that when liver cells accumulate fat, they change which genes are turned up or down. Some of the genes that were turned up are used by cells to make lipid droplets.

I also found that a lot of genes were turned down. A lot of these genes provide instructions to make proteins for structures called mitochondria. Mitochondria are found in all cells in the body and use sugars and fats to generate most of the energy that the body needs, leading to them being called the “powerhouses of the cell”. When people develop NAFLD, mitochondria do not work as well and cannot use sugars and fats properly, but it is unclear how this is linked to disease. I used the cell model of NAFLD to study whether mitochondria in liver cells with excess fat stop working properly. I found that mitochondria in these cells use sugar differently to healthy cells and this leads them to make more of a chemical called fumarate. I found that when mitochondria make too much fumarate, this leads them to make bigger lipid droplets.

Fumarate is also important because it can directly affect the DNA in the cell. Genes sometimes have extra chemicals (methyl groups) attached to them at specific locations. The methyl group (5mC) can be converted to a hydroxymethyl (5hmC) group by a group of proteins called the ten-eleven translocation (TET) enzymes. In specific circumstances, these chemical groups can be used to instruct the gene to make a protein, stop making a protein, or to make more or less of it. Fumarate can prevent the TET enzymes from working, which also interferes with the conversion of 5mC to 5hmC. However, I did not find evidence that the TET enzymes are responsible for turning up genes related to making proteins associated with lipid droplets. This indicates that, in NAFLD, fumarate affects proteins other than the TET enzymes, which can also influence how genes are turned on, off, up or down. This information might be useful for future studies that try to develop treatments for NAFLD.

Publications

Publication 1	**Sinton MC**, Hay DC, Drake AJ (2019) Metabolic control of gene transcription in non-alcoholic fatty liver disease: the role of the epigenome, Clinical Epigenetics 11, 104
Author's contributions – general	**MCS** conceived and wrote the manuscript with AJD. DCH edited the manuscript and provided advice. All authors read and approved the final version of the manuscript.
Confirmation	Date: 24th July 2020 Signature: Print full name: Prof. Mandy J. Drake
Publication 2	**Sinton MC**, Meseguer Ripolles J, Lucendo Villarin B, Lyall MJ, Carter RN, Morton NM, Hay DC, Drake AJ (2020) Nonalcoholic fatty liver disease is associated with decreased hepatocyte mitochondrial respiration, but not mitochondrial number, bioRxiv, doi: 10.1101/2020.03.10.985200
Author's contributions – general	**MCS**, JMR, BLV, MJL, RNC and NMM performed experiments. **MCS**, DCH and AJD conceived the experiments. **MCS**, NMM, DCH and AJD wrote the paper. All authors contributed to drafts of the paper and approved the final version.
Candidate's contributions – specific	1) Stem cell maintenance and differentiation; 2) optimisation of Seahorse protocol; 3) RNA purification and RT-qPCR; 4) western blotting; 5) data analysis
Confirmation	Date: 24th July 2020 Signature: Print full name: Prof. Mandy J. Drake
Publication 3	**Sinton MC**, Lucendo Villarin B, Meseguer-Ripolles J, Wernig-Zorc S, Thomson JP, Walker PD, Thakker A, Lavery GG, Ludwig C, Tennnant DA, Hay DC, Drake AJ (2020) Macrovesicular steatosis in nonalcoholic fatty liver disease is a consequence of purine nucleotide cycle driven fumarate accumulation, bioRxiv, doi: 10.1101/2020.06.23.166728

Author's contributions – general	**MCS** was involved in conceptualisation, methodology, validation, formal analysis, investigation, data curation, writing (original draft preparation, reviewing and editing) and visualisation. BLV, JMR, PDW and AT were involved in methodology, validation, formal analysis, investigation and writing (reviewing and editing). SW-Z and JPT were involved in software, formal analyses, data curation and writing (reviewing and editing). CL and DAT were involved in conceptualisation, methodology, validation, formal analysis, investigation, data curation, provision of resources, writing (reviewing and editing), project administration, funding acquisition and supervision. DCH and AJD were involved in conceptualisation, methodology, validation, formal analysis, provision of resources, writing (reviewing and editing), project administration, funding acquisition and supervision.
Candidate's contributions – specific	1) RNA purification and QC; 2) isotopic tracing and metabolite harvesting; 3) inhibitor treatments; 4) GC-MS/NMR data interpretation; 5) DNA purification, genome amplification & hmeDIP-seq library preparation; 6) hmeDIP-seq sequencing analysis
Confirmation	Date: 24th July 2020 Signature: Print full name: Prof. Mandy J. Drake

Acknowledgements

I want to thank a number of people, without whom this thesis would not have been possible. I want to start by saying the biggest thank you to my supervisor, Professor Mandy Drake. Mandy is the most supportive supervisor imaginable, lifting me up when I was down, giving me space to explore and grow, and providing so many opportunities in my time with her. I will always value being given compassion and encouragement that I received from Mandy. It is because of Mandy's style of mentorship and supervision that I never questioned my decision to undertake a PhD and have always known that I picked the right lab.

I also want to thank members of Team Drake, particularly Marcus and Bonnie. I thank Marcus for putting the time and effort into developing the model that I worked with throughout my PhD. Bonnie, thank you so much for the kindness and patience you showed me when training me, and for listening to my hare-brained ideas. I am grateful to my second supervisor, Professor Dave Hay, for all of the support and feedback he provided throughout my PhD and for letting me use lab space and resources. Within Dave's lab, I owe a massive debt of gratitude to Balta and Jose. Jose, thank you for training me in stem cell culture and differentiation, even when I drove you mad with endless questions and knocking things over in hoods, as well as for all of the help over weekends. Balta, thank you for all of the help with differentiations and giving me advice and support. I really couldn't have managed to do half of what I did without you two. Thank you to Professor Nik Morton and Dr Rod Carter for their help in designing and running the Seahorse mitochondrial stress tests, and for fruitful conversations. For help with my DIP-seq analysis, I need to thank Sara Wernig-Zorc and John Thomson, for all of their advice. Thank you also to Dr Will Cawthorn and Dr James Minchin for hosting me in their labs for rotation projects. My time in your labs made it incredibly difficult to choose which project I wanted to pursue for a PhD.

At the University of Birmingham, I want to thank Daniel Tennant, Christian Ludwig, Alpesh Thakker, Paul Walker and Gareth Lavery. I thoroughly enjoyed my discussions

with all of you, learning about mitochondrial metabolism, and the technologies used to examine it. Of course, I have to say thank you for providing protocols and for running the GC-MS and NMR spectroscopy. The data from those experiments made my project really exciting and really spurred me on when I was struggling to keep motivated.

Beyond the university, the biggest thanks go to my husband, Juan, for believing in me when I didn't believe in myself, and for encouraging me to be the best version of myself. Before returning to academia, I doubted whether I could make it, and without your support I might never have taken the first steps back onto this path. You are my everything, and I will be forever grateful that you came into my life. Our dog, Igor, has been the best companion throughout this journey, distracting me and making me laugh and smile. I also need to thank my family, particularly my parents. My career has taken may twists and turns and, whilst they weren't always certain of my decisions, my parents always supported me in any way that they could. I also have to thank so many friends for their support throughout this process, but in particular Lorna, Sara, Shiny, as well as all of my friends at CVS. I also want to thank those people who gave me my love of biology, and particularly metabolism, before I started the PhD, including my sixth form college tutor Pamela Cowey, and my undergraduate tutors David Wright and Helen Page.

Lastly, I want to thank everyone at CVS, for all of the help and support over the past 4 years, which has made my PhD journey so very special. I will always remember this period of my life with fondness and gratitude.

Abbreviations

5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
ACACA	Acetyl-CoA carboxylase
AICAR	5-Aminoimidazole-4-carboxamide-1-β-D-ribofuranosyl 5′-monophosphate
αKG	Alpha-ketoglutarate
AKR	5β-reductase
ALB	Albumin
AMPK	5’ AMP-activated kinase
AOA	O-(Carboxymethyl)hydroxylamine hemihydrochloride
ATP	Adenosine triphosphate
ATP5/6	ATPase subunits 5/6
BER	Base excision repair
CASP1	Caspase 1
CGI	CpG island
ChREBP	Carbohydrate-responsive element-binding protein
COL6A1	Collagen type VI alpha 1
COL12A1	Collagen type XII alpha 1
COX	Cyclooxygenase
CP	Ceruloplasmin
CpG	Cytosine-phosphate-guanine
CPT1	Carnitine palmitoyltransferase 1
CS	Citrate synthase
CTNNB1	Catenin Beta 1
CYP7A1	Cholesterol 7 alpha-hydroxylase
DAVID	Database for Annotation, Visualization, and Integrated Discovery
DEG	Differential gene expression
DGAT2	Diacylglycerol acetyltransferase 2
DHR	Differentially hydroxymethylated region
DIP	DNA immunoprecipitation
DNA	Deoxyribonucleic acid

DNL	*De novo* lipogenesis
DNMT	DNA methyltransferase
DOC2A	Double C2-like domain-containing protein alpha
ECM	Extracellular matrix
EPHX3	Epoxide hydrolase 3
ER	Endoplasmic reticulum
EROB1	Endoplasmic reticulum oxidoreductase 1β
ESC	Embryonic stem cell
ETC	Electron transport chain
FASN	Fatty acid synthase
FCCP	Carbonyl-cyanide-4-(trifluoromethoxy) phenylhydrazone
FFA	Free fatty acid
FGF19	Fibroblast growth factor 19
FH	Fumarate hydratase
FOXO1	Fork-head box protein
FXR	Farnesoid X receptor
G6PD	Glucose-6-phosphate dehydrogenase
GC-MS	Gas chromatography-mass spectrometry
GLUT1	Glucose transporter 1
GLUT4	Glucose transporter 4
GO	Gene ontology
GSK3β	Glycogen synthase kinase-3β
H3	Histone 3
H4	Histone 4
HAMP	Hepcidin
HBB	Beta-globulin
HBP	Hexosamine biosynthetic pathway
HCC	Hepatocellular carcinoma
HDAC8	Histone deacetylase 8
HIF	Hypoxia inducible factor
HLC	Hepatocyte-like cells
HMGCS1/2	3-Hydroxy-3-Methylglutaryl-CoA Synthase ½
HMW	High molecular weight
HNF4A	Hepatocyte nuclear factor 4 alpha

hPSC	Human pluripotent stem cell
HSC	Hepatic stellate cell
IDH	Isocitrate dehydrogenase
IL-6	Interleukin 6
IR	Insulin resistance
JHDM	Jumonji domain-containing histone demethylase
KEGG	Kyoto Encyclopedia of Genes and Genomes
KC	Kupffer cell
LDH	Lactate dehydrogenase
LPO	Lactate, pyruvate, octanoic acid
MAS	Malate aspartate shuttle
MBD	Methyl-CpG-binding domain
MDD	Methionine-deficient diet
MDH	Malate dehydrogenase
MTP	Microsomal transfer protein
NAFLD	Non-alcoholic fatty liver disease
NANOG	Homeobox Transcription Factor Nanog
NASH	Non-alcoholic steatohepatitis
ND1/2	NADH dehydrogenase ½
NDUF	NADH:ubiquinone oxidoreductase core subunits
NF-κB	Nuclear factor kappa B
NK cell	Natural killer cell
NLRP3	NOD-, LRR- and pyrin domain-containing protein 3
NMR	Nuclear magnetic resonance
NR0B2	Nuclear Receptor Subfamily 0 Group B Member 2
OAA	Oxaloacetate
OCR	Oxygen consumption rate
OGDHL	2-oxoglutarate dehydrogenase-like
OGT	*O*-linked N-acetylglucosamine transferase
ORM1/2	Orosomucoid 1/2
PHD	Prolyl-4-hydroxylase
PC	Pyruvate carboxylase
PCA	Principal component analysis
PDH	Pyruvate dehydrogenase

PHD	Prolyl-4-hydroxylase
PINK1	PTEN-induced putative kinase
PLIN2/3/4/5	Perilipin 2/3/4/5
PNC	Purine nucleotide cycle
PNPLA3	Patatin-like phospholipase domain-containing protein 3
PPARα	Peroxisome proliferator-activated receptor alpha
PPARγ	peroxisome proliferator-activated receptor γ
PPARGC1A	PPARγ coactivator 1α
PRC	Polycomb recessive complex
PRKN	Parkin
PRR	Pattern recognition receptor
PTM	Post-translational modification
PYGO1	Pygopus Family PHD Finger 1
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RPL7	60S ribosomal protein L7
RPL11	60S ribosomal protein L11
RPMI	Roswell Park Memorial Institute
RPS19	40S ribosomal subunit S19
RPS24	40S ribosomal subunit S24
SAA4	Serum Amyloid A4
SDHA	Succinate dehydrogenase subunit A
SLC2A4	Solute carrier family 2 member 4
SNP	Single nucleotide polymorphism
SOCS3	Suppressor of cytokine signalling 3
SOD1/2	Superoxide dismutase 1/2
SRB	Sulforhodamine B
SREBP-1c	Sterol regulatory element-binding protein
SUCLG1	Succinate-CoA ligase GDP/ADP-forming subunit alpha
SULT	Sulfotransferase
T2D	Type 2 diabetes
TCA	Tricarboxylic acid
TES	Transcriptional end site
TET	Ten-eleven translocation dioxygenases

TFR2	Transferrin Receptor 2
TG	Triglyceride
TGF-β	Transforming growth factor beta
TGH	Triglyceride hydroxylase
Th1	Type 1 helper cell
Th2	Type 2 helper cell
TLR4	Toll-like receptor 4
TMEM88	Transmembrane Protein 88
TNF-α	Tumour necrosis factor alpha
TSS	Transcriptional start site
UCP2	Uncoupling protein 2
UGT	UDP-glucuronosyltransferases
UPLC	Ultra-performance liquid chromatography
UPR	Unfolded protein response
VLDL	Very low-density lipoprotein

Chapter 1: Introduction

1.1 The pathogenesis of non-alcoholic fatty liver disease (NAFLD)

1.1.1 Background of NAFLD

In 2016 The Lancet published a survey of 19.2 million people, which revealed that, for the first time in history, the number of overweight people around the world outnumbers those who are malnourished [1]. This global obesity epidemic contributes to an approximated 2.8 million deaths per year [2], and is associated with the development of numerous diseases, including hypertension, type 2 diabetes (T2D), coronary heart disease, and NAFLD [3]. In line with the growing obesity epidemic, NAFLD has become the most common form of hepatic disease globally [4]. Whilst the association between obesity and NAFLD has been recognised since the 1960s [5], the functional link between these two disorders was not identified until the late 1970s and early 1980s [6–9]. In 1979, histology analysis revealed a range of distinct subtypes in the livers of obese patients, from simple lipid accumulation (benign steatosis), to inflammatory hepatitis, fibrosis and cirrhosis (**Figure 1.1**) [6]. It was later revealed that the development of NAFLD is an independent risk factor for the development of hepatocellular carcinoma (HCC), the second most common cause of cancer-related death [10]. Together, these findings demonstrated that NAFLD has the potential to progress to life-threatening disease. At present, the only way to reverse the development of benign steatosis or non-alcoholic steatohepatitis (NASH) is through diet and exercise [11]. As the obesity epidemic continues, the number of people living with NAFLD will continue to grow, increasing the economic burden on healthcare systems globally [12], making it critical that new strategies for the treatment of the disease are developed.

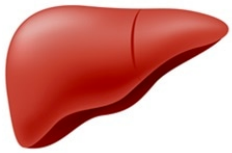
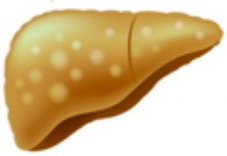
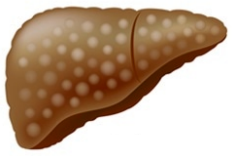
Pathology		Risk Factors
Healthy liver		
		
Benign steatosis ~90% of obese adults		Obesity Insulin resistance Dyslipidaemia Family history
		
NASH 20% of NAFLD patients		Obesity T2DM Metabolic syndrome
		
Cirrhosis 11% of NASH patients		Unknown

Figure 1.1 Progression of NAFLD and associated risk factors in adult populations. Arrows indicate direction of disease progression and whether disease is reversible following intervention. Data are from Marengo *et al* (2016) [13].

1.1.2 Histopathology of NAFLD

As stated above, NAFLD comprises a spectrum of pathologies, from benign hepatic steatosis, to inflammatory NASH, liver cirrhosis, and HCC [14]. Histological analysis of NAFLD remains the gold standard for the diagnosis of NAFLD, whether this is benign steatosis or NASH [15]. Since histological analysis requires the collection of

biopsies, which is invasive and carries a risk, there is a need to develop new methods of diagnosis and study [16].

1.1.2.1 Benign Steatosis

Benign steatosis is characterised by hepatic fat content greater than 5% of total liver weight [17]. In adults, steatosis is normally characterised by macrovesicular lipid droplets, referring to a single large cytoplasmic lipid droplet, which can displace the nucleus to the periphery of the cell [18]. Less commonly, steatotic hepatocytes may contain multiple smaller lipid droplets (microvesicular steatosis), which do not displace the nucleus (**Figure 1.2A**) [18]. These lipid droplets are composed of multiple lipid species, including triglycerides (TGs), ceramides, diacylglycerides, cholesterol esters and free fatty acids (FFAs) [19]. In adults, steatotic cells are generally scattered around the small branches of the hepatic vein (acinar zone 3) [16]. This is in contrast to obese adolescents, where steatotic hepatocytes are more likely to be located in the periportal region [20]. In more extensive cases, distribution of steatotic hepatocytes can become random, occurring throughout the liver [21].

1.1.2.2 NASH

Approximately 37% of adult patients with benign steatosis will progress to develop NASH, which is strongly associated with hepatic inflammation, insulin resistance (IR), and T2D [22]. The development of NASH is characterised by hepatic inflammation and hepatocyte ballooning (**Figure 1.2B**), leading to tissue injury and remodelling [23]. As inflammation develops, the liver is infiltrated by a range of immune cells, including macrophages, natural killer (NK) cells, eosinophils, and type 1 helper T (Th1) cells [24]. Eosinophils deposit protein aggregates, termed Mallory-Denk bodies (**Figure 1.2C**), which are used in histology to identify NASH [25]. NASH is also linked to the development of megamitochondria (**Figure 1.2D**), but the contribution of these enlarged organelles to NASH pathology remains unclear [26].

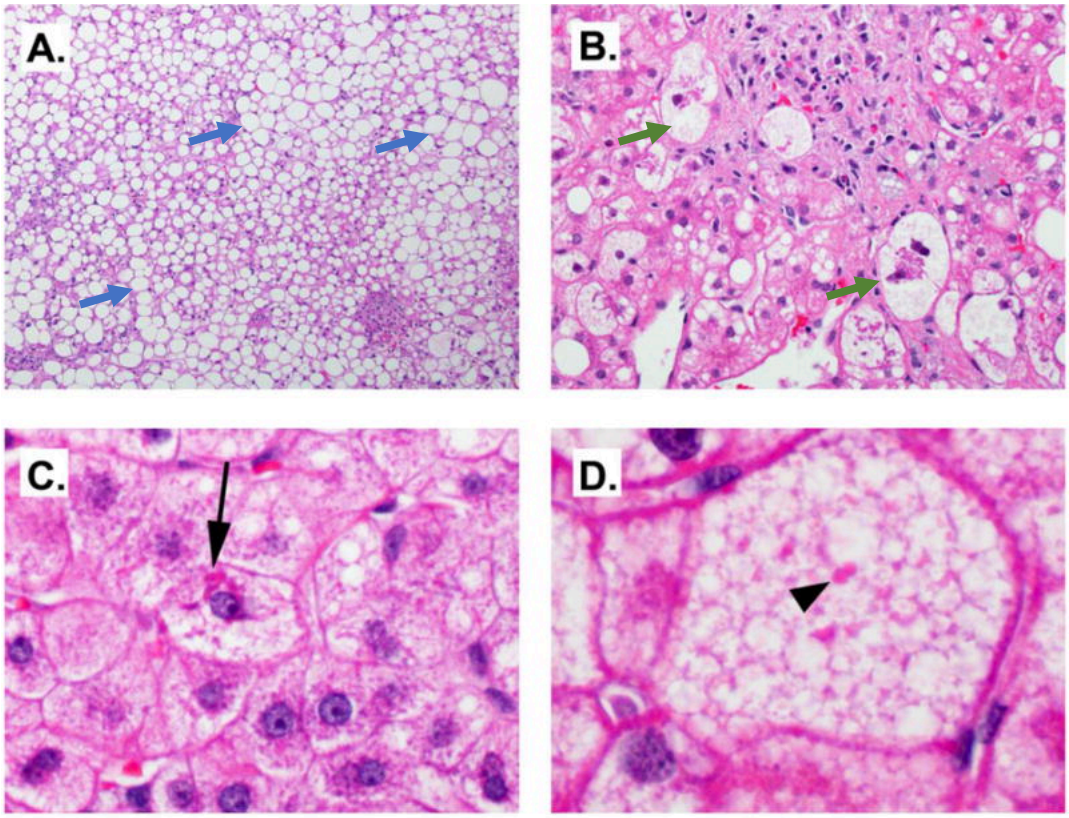


Figure 1.2 Histopathological analysis of NAFLD and NASH. (**A**) Steatotic hepatocytes are characterised by macrovesicular lipid droplet accumulation (*blue arrows*). (**B**) NASH is characterised by development of hepatocyte ballooning (*green arrows*). (**C**) Mallory-Denk bodies (*black arrow*) are a sign of hepatocyte injury. (**D**) Development of megamitochondria (*black arrowhead*) is associated with NASH. Figure adapted from Brown & Kleiner, 2016 [25].

1.1.2.3 Cirrhosis

Cirrhosis is the endpoint of the majority of chronic liver diseases and is characterised by the extensive deposition of collagen by fibroblasts, leading to a loss of tissue function [13]. Approximately 11% of patients with NASH-associated fibrosis will go on to develop cirrhosis [13]. This causes the irreversible loss of the liver parenchyma, limiting the ability of the organ to synthesise coagulation factors and albumin, as well as impairing blood flow and detoxification capacity. The development of this

phenotype limits the ability of the liver to function effectively, preventing it from excreting bilirubin, and resulting in jaundice and eventual organ failure [27].

1.1.3 Progression of NAFLD to NASH and development of fibrosis

1.1.3.1 From the two-hit to multiple hit model of NAFLD

A major hurdle in the development of therapeutics for NAFLD is that the cause is multi-factorial [28]. In 1998, Day and James proposed the two-hit model, whereby benign steatosis represents the first “hit”, characterised by the accumulation of intracellular TGs. The second “hit” is characterised by the development of an inflammatory phenotype, associated with the production of inflammatory cytokines, oxidative stress, and mitochondrial dysfunction, triggering the development of NASH [29]. In this thesis, mitochondrial dysfunction is defined as alterations in mitochondrial function associated with pathogenesis. This has developed into a multiple-hit model, which considers numerous other factors driving pathology, including diet and genetic predisposition [30]. There is increasing evidence that dysfunctional mitochondrial metabolism may impact on epigenetic modifications of DNA and chromatin, and may also play a role in the development of this disease [31–33]. In particular, studies in humans and mice indicate that mitochondrial tricarboxylic acid (TCA) cycle disruption plays a central role in the pathogenesis of NAFLD [34,35]. These points will be discussed later in the introduction.

1.1.3.2 The role of lipid species in progression from benign steatosis to NASH

The development of lipid droplets during NAFLD pathogenesis may represent a protective mechanism employed by hepatocytes that are overloaded with fatty acids (FAs) [36]. This is supported by mouse studies of diacylglycerol acetyltransferase 2 (DGAT2), which catalyses the final step in TG synthesis. Knockdown of DGAT2 reduced accumulation of hepatic TGs but was associated with exacerbation of

hepatocellular injury [37]. If sequestration of TGs is protective, then this suggests that the generation of non-TG lipids plays a role in hepatocellular injury, progression to NASH, and development of fibrosis. The harmful effect of non-TG lipids is termed lipotoxicity, and in hepatocytes this is understood to contribute to the conversion of benign steatosis to NASH [36]. These non-TG lipids include diacylglycerol, acylcarnitines, FAs, cholesterol and ceramides, all of which are known to induce hepatocellular injury [38] and are also elevated in NASH patients [39,40]. Induction of hepatocellular injury by these non-TG lipids occurs through a number of pathways. For example, high levels of FA clearance by microsomal, peroxisomal and beta-oxidation generates reactive oxygen species (ROS), causing oxidative stress and cellular injury [41]. High levels of intracellular FAs and cholesterol also induce endoplasmic reticulum (ER) stress and activation of the unfolded protein response (UPR) [42,43], a known contributor to NASH [44,45]. This is thought to impact on mitochondrial function too due to disruption of Ca^{2+} transfer between the mitochondria and ER [46]. The UPR is also known to induce ceramide synthesis [47], with high levels of ceramides inducing excessive ROS generation and apoptosis in steatotic hepatocytes [40,48]. Taken together, these combined findings support the protective role of TG sequestration in lipid droplets and the importance of non-TG lipids in the transition from benign steatosis to NASH.

1.1.3.3 Inflammation

Progression of benign steatosis to NASH is accompanied by infiltration of the hepatic parenchyma by innate and adaptive immune cells [49]. Lipotoxicity-induced hepatocellular injury and apoptosis triggers the release of danger-associated molecular pattern (DAMP) molecules, which bind to pattern recognition receptors (PRRs), including Toll-like receptor 4 (TLR4) [50]. Binding of DAMPs to TLR4 on Kupffer cells (KCs) activates nuclear factor kappa B (NF-κB) signalling, leading to secretion of cytokines such as interleukin 6 (IL-6) and tumour necrosis factor alpha (TNF-α). Whole-body knockout of TLR4 in mice attenuates the development of NASH, indicating its central role in the transition from NAFLD to NASH [50]. Activated KCs begin secreting transforming growth factor beta (TGF-β) and recruit other immune

cells, including neutrophils and natural killer cells, which infiltrate the liver parenchyma and promote inflammation [51].

1.1.3.4 Fibrosis

Hepatic inflammation and secretion of TGF-β leads to activation of liver-resident hepatic stellate cells (HSCs), which are normally quiescent, leading them to transdifferentiate to a myofibroblast phenotype [52]. Following transdifferentiation, HSCs become proliferative and chemotactic, and begin secreting extracellular matrix (ECM) proteins [53], including type I and type III collagens, laminin, and proteoglycans, leading to development of fibrosis [54–56]. Although hepatic fibrosis is not used as a diagnostic feature of NASH, its presence is a predictor of mortality and whether patients will require a liver transplant [57]. Where NASH-associated inflammation is not resolved, this can lead to excessive fibrosis and, ultimately, cirrhosis [58], but it is unclear why some patients develop this feature of NAFLD and others do not.

1.2 Molecular basis of NAFLD

1.2.1 The thrifty gene hypothesis

NAFLD comprises part of the metabolic syndrome, a cluster of interconnected conditions including obesity and IR, which increase the risk of developing cardiovascular disease and T2D [59]. In 1962, the “thrifty gene hypothesis” was formulated, and posited an evolutionary selection of genes that contribute to increased lipid storage, to use as an energy source in times of food scarcity [60]. In modern society, where humans have access to abundant quantities of food, the prevalence of these genes increases the risk of developing disease, rather than being protective and providing an evolutionary advantage. Whilst there are undoubtedly genes that contribute to the development of components of the metabolic syndrome, there is some dispute over whether these genes provide an evolutionary advantage in times of food scarcity [61]. However, identification of “thrifty” genes may aid in the

search for therapeutic targets for the treatment of components of the metabolic syndrome.

1.2.2 Genetics of NAFLD

There have been a number of genome-wide association studies (GWAS) identifying alleles associated with increased risk of developing NAFLD [62–65]. These studies identified multiple single nucleotide polymorphisms (SNPs) associated with NAFLD, including PNPLA3 [62–65]. PNPLA3 is a membrane-associated TG hydrolase, involved in the packaging of TGs into very low-density lipoproteins (VLDLs) prior to export [66]. Pirazzi *et al* (2012) studied a mixed ethnicity population with increased hepatic lipid levels and found an isoleucine to methionine amino acid substitution that was independently associated with the NAFLD phenotype [66]. Further studies demonstrated that the I148M SNP in PNPLA3 is strongly associated with more extensive steatosis, lobular inflammation and development of fibrosis [67], but not with hepatic IR [68]. Furthermore, this SNP was associated with reduced VLDL export in steatotic livers [66,69–71]. There are also studies indicating that SNPs in genes associated with one-carbon metabolism may also correlate with severity of steatosis in NAFLD [72]. Unlike PNPLA3, this was observed in a single study, but as perturbation of one-carbon metabolism is associated with NAFLD [73], the association may be relevant. Recently, a GWAS of a European cohort identified the Pygopus family PHD finger 1 (PYGO1) gene as a novel modifier of steatosis, suggesting that the Wnt signalling pathway is also relevant to NAFLD pathogenesis [74].

1.2.3 Insulin resistance

IR is a hallmark of NAFLD and is tightly linked to the pathogenesis of the disease when associated with obesity [75]. Although GWAS studies have shown no association between any SNP and hepatic IR, NAFLD is associated both with hepatic and systemic IR [68]. However, the molecular basis of this association is not well understood. What is known, however, is that in obesity, the excess accumulation of lipids in visceral adipose tissue results in an increased release of FFAs into the

circulation [76]. In the steady state, stimulation of hepatic insulin receptors activates the PI3K/AKT pathway, inducing phosphorylation of the fork-head box protein FOXO1 transcription factor [77] and glycogen synthase kinase-3β (GSK3β) [78]. Phosphorylation of FOXO1 prevents it from entering the nucleus [77]. However, disruption of the insulin signalling pathway can prevent the phosphorylation of FOXO1, allowing it to translocate to the nucleus, where it stimulates expression of genes associated with gluconeogenesis and *de novo* lipogenesis [79]. This leads to hepatic intracellular lipid accumulation, and development of large lipid droplets within the cytoplasm [16]. Lipid droplet accumulation is associated with increased expression of perilipin 2 (PLIN2), which regulates GSK3β activation through direct interaction with the enzyme [80]. This indicates that as well as mediating IR, GSK3β is responsive to intracellular lipid accumulation, and may be a key link between lipid accumulation and hepatic IR. Involvement of GSK3β in NAFLD was further demonstrated using a GSK3β knockout mouse, which failed to develop hepatic steatosis when fed a high fat diet, due to altered activity of adiponectin [81].

1.2.4 Mechanisms of lipid accumulation in NAFLD

The classic feature of NAFLD is the accumulation of intracellular lipid droplets within hepatocytes, which is associated with increased mitochondrial β-oxidation, and reactive oxygen species (ROS) generation [82–84]. This occurs as a result of imbalances in lipid uptake and synthesis compared to lipid export [85]. Whilst lipid uptake by hepatocytes is most likely a normal physiological process, the pathological accumulation of lipid droplets in steatotic livers is a consequence of chronic energy imbalance [85].

1.2.4.1 Adipose tissue dysfunction causes excess fatty acid delivery to hepatocytes

In normal weight adults, in the fed state, insulin secretion inhibits hydrolysis of TGs, but where patients are insulin resistant this causes a state of chronic inflammation. In turn, this leads to excess lipolysis within adipocytes, and subsequent release of FFAs

into the serum, which are taken up by the liver [85]. Patients with NAFLD display elevated serum FFA levels, induced by increased adipose tissue lipolysis [82,86]. Studies estimate that ~60% of TGs in steatotic lipids are derived from these serum FAs, with a minority generated by *de novo* lipogenesis (DNL) [83]. In the same study, radiolabelling of non-esterified fatty acids (NEFAs) revealed a small contribution of lipids in steatotic livers from dietary fats. Elevation of the fatty acid (FA) transporter CD36 was also identified in human steatotic livers, indicating an adaptation to enable increased FA sequestration from serum [87,88].

As well as its role in IR and release of FFAs, adipose tissue also promotes lipid accumulation through secretion of adipose tissue hormones (adipokines) and inflammatory mediators. The adipokine adiponectin is secreted in multiple isoforms of varying molecular weight. Low level adipocyte secretion of the high molecular weight (HMW) isoform of adiponectin is observed in patients with IR, obesity, and NAFLD [89,90]. During fasting, high levels of circulating HMW adiponectin bind to hepatocyte receptors, increasing peroxisome proliferator-activated receptor alpha (PPARα) and 5' AMP-activated kinase (AMPK) activity [91]. In turn, increased PPARα and AMPK activity inhibit lipogenesis, promoting FA oxidation, providing substrates for the TCA cycle. In contrast to obesogenic conditions, this process in fasting promotes lipid clearance [91]. NAFLD and obesity are also associated with high levels of circulating IL-6 and TNFα. IL-6 is a pleiotropic cytokine, which has both inflammatory and anti-inflammatory roles [92]. In mouse studies of obesity, IL-6 appears to promote lipolysis and protect against obesity [93]. However, in contradiction to this, increased levels of IL-6 are associated with the development of IR, which, in turn, is associated with the development of obesity and NAFLD [94], demonstrating the complexity of unpicking the role of this cytokine. TNFα is a pro-inflammatory cytokine, predominantly secreted by macrophages. In morbidly obese humans with NAFLD, circulating TNFα levels are elevated [95], suggesting that it plays a role in obesity-related systemic inflammation. Mouse studies show that inhibition of TNFα signalling protects against the development of steatosis, indicating that it plays a key role in the pathogenesis of these conditions [96]. Furthermore, both IL-6 and TNFα suppress insulin signalling by promoting the upregulation of suppressor of cytokine signalling 3 (SOCS3), further

contributing to lipid release from adipocytes [97]. The accumulation of hepatic lipids is associated with the development of mitochondrial dysfunction, but the mechanism linking these two disease features is unclear [98].

1.2.4.2 Contribution of DNL to intracellular lipid accumulation in hepatocytes

Whilst DNL contributes less to the TG pool than FAs, it is still a major source of lipids contributing to NAFLD [83]. This process synthesises FAs from acetyl-CoA subunits. These subunits are converted to malonyl-CoA by acetyl-CoA carboxylase (ACACA) [99]. Elevation of liver malonyl-CoA inhibits the rate-limiting enzyme of beta-oxidation, carnitine palmitoyltransferase 1 (CPT1), inhibiting, in turn, FA oxidation [100]. Fatty acid synthase (FASN) then converts malonyl-CoA to the long-chain saturated fatty acid palmitate [99]. This process is tightly regulated at the transcriptional level, with glucose and insulin activating the carbohydrate-responsive element-binding protein (ChREBP) and sterol regulatory element-binding protein (SREBP-1c), respectively [101,102]. In response to glucose and insulin, ChREBP and SREBP-1c are able to modulate the expression of genes required for DNL [103,104]. The overexpression of SREBP-1c, FASN and ACACA has been observed in NAFLD patients, further supporting that DNL plays a role in human disease [105].

1.2.4.3 Impaired lipid export

In addition to increased levels of TG synthesis, NAFLD is associated with impaired export of lipids. Under homeostatic conditions, lipid export functions to deliver TGs and cholesterol to other tissues in the body, including cardiac and skeletal muscle [106]. These are transported in the form of VLDLs, which are lipoprotein particles composed of a lipid bilayer containing TGs and cholesterol esters, and stabilised by a single apolipoprotein B-100 molecule [106]. Assembly of VLDLs occurs in the lumen of the ER, before transport to the Golgi body. Following modification in the Golgi body, VLDLs are trafficked to the cell membrane and exported from the cell [107]. Assembly in the ER is facilitated by a range of hydrolase enzymes, including triglyceride

hydroxylase (TGH), microsomal transfer protein (MTP), carboxylesterase enzymes, and PNPLA3. As discussed previously in this chapter, PNPLA3 was identified in a number of GWAS studies as a high-risk allele for developing hepatic steatosis [63–65]. Whilst the export process of VLDLs continues to function in the steatotic liver, excretion of these particles increases, plateauing once hepatic fat content reaches ~10% [82]. This indicates that under steatotic conditions, the liver continues to export lipids as far as possible but is unable to do so once this limit is reached. Increased export of VLDLs increases ER stress, suggesting that increased assembly and export of these proteins may also impact on mitochondrial function [46].

1.3 The role of mitochondria in NAFLD

Mitochondria are organelles found in eukaryotic cells and are often termed the "powerhouse of the cell", due to their central role in regulating energy homeostasis [108]. Further to this, they are also key regulators of cell death [108]. Multiple studies suggest that the pathology of NAFLD is linked to impairment of mitochondrial function [34,109–111]. In rat models, mitochondrial dysfunction precedes the development of hepatic steatosis, suggesting that disease pathology and progression are linked to impairment of mitochondrial function [32]. NAFLD is linked to structural changes in mitochondrial cristae and the development of crystalline inclusion bodies [112]. As electron transport chain (ETC) components are embedded in the cristae, structural changes may have a physical impact on respiration [113]. Similarly, the growth of inclusion bodies may also impair oxidative phosphorylation and lead to increased ROS production [112]. This is supported by observations of NAFLD-associated hepatic oxidative stress and altered levels of fatty acid beta oxidation [112,113]. Furthermore, in humans, once hepatic lipid content exceeds 10%, lipid export is compromised [82]. In tandem with altered fatty acid beta oxidation, this imbalance may lead to the accumulation of intracellular lipids observed in NAFLD.

1.3.1 Mitochondrial quality control

Due to their central role in energy metabolism and cell death, it is important that damaged or dysfunctional mitochondria are regularly removed [108]. This occurs via a quality control process termed mitophagy, whereby damaged organelles are enveloped by a double-membraned structure called the autophagosome, and trafficked to the lysosome for degradation [114]. Mitophagy is regulated, in part, by the PINK1-PRKN signalling pathway, whereby the PTEN-induced putative kinase (PINK1) phosphorylates the E3 ubiquitin ligase Parkin (PRKN) [115]. PINK1 is located on the outer mitochondrial membrane, and is regulated by changes in membrane potential [116]. When this potential is disrupted in damaged mitochondria, it stabilises PINK1, which recruits PRKN, enabling the ligase to ubiquitinate the organelle and target it for degradation [116]. This signalling pathway also impacts on NAFLD, with mouse models showing that hepatic knockout of *Prkn* increased steatosis and reduced mitochondrial respiratory capacity in hepatocytes [117]. These mice also displayed reduced insulin sensitivity, despite no changes in weight gain following a high fat diet. Although it is not currently known whether this applies to human disease, it strongly indicates that mitochondrial integrity and function is central to maintaining hepatic homeostasis and protecting against the development of NAFLD.

1.3.2 The TCA cycle and electron transport chain

The TCA cycle occurs predominantly in the mitochondrial matrix, and is a central pathway for the metabolism of amino acids, fatty acids, and carbohydrates, to generate cellular energy [118]. The cycle utilises the substrate acetyl coenzyme A (acetyl-CoA), which is generated by β-oxidation, as well as amino acid degradation and glycolysis *via* pyruvate. Acetyl-CoA is oxidised through a series of steps (**Figure 1.3**), producing nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide ($FADH_2$), which subsequently transfer electrons to the electron transport chain (ETC), to fuel the production of adenosine triphosphate (ATP) (**Figure 1.4**). Acetyl-CoA is generated by the oxidation of pyruvate, by pyruvate dehydrogenase (PDH) [119]. Pyruvate can also be carboxylated by pyruvate carboxylase (PC) to form oxaloacetate (OAA). Acetyl-CoA and OAA then undergo a condensation reaction to

form citrate, which can then progress through the TCA cycle, forming multiple intermediates, including alpha-ketoglutarate (αKG), succinate, fumarate, and malate. In addition to acetyl-CoA, the TCA cycle can also be fuelled by the deamination of glutamate by glutamate dehydrogenase, to produce αKG [120]. Increased influx of substrates into the TCA cycle can also increase cycle flux, as well as the abundance of the intermediates αKG, succinate, and fumarate [121].

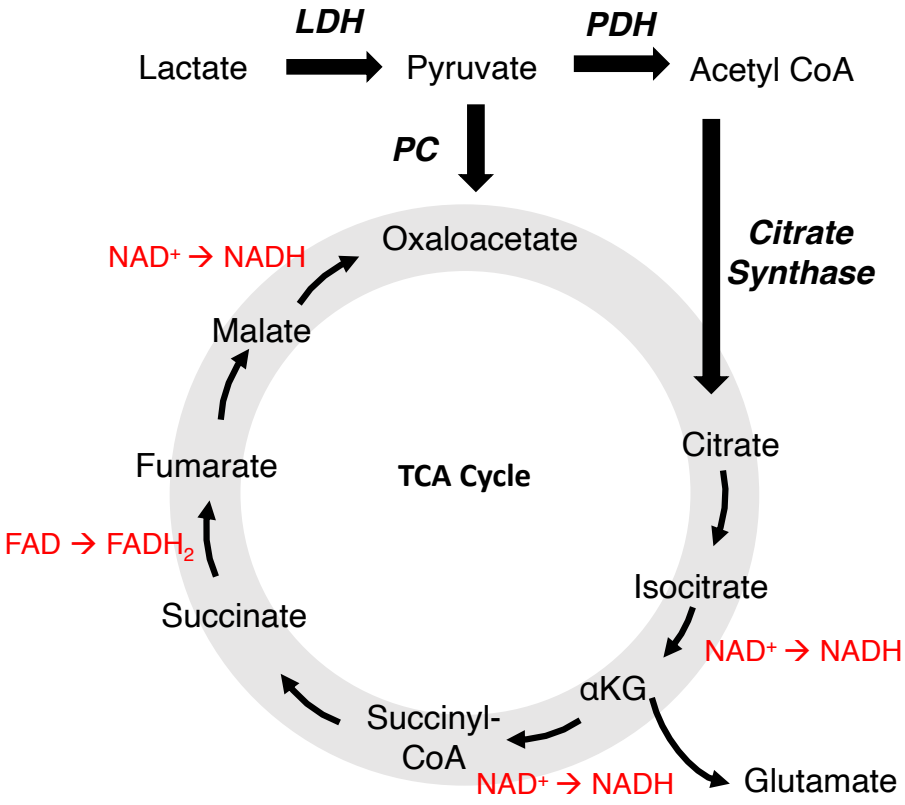


Figure 1.3. Schematic of the TCA cycle. Arrows indicate the direction of metabolite conversion under homeostatic conditions. Red text denotes processes generating substrates for the electron transport chain. LDH = lactate dehydrogenase; PDH = pyruvate dehydrogenase; PC = pyruvate carboxylase; aKG = alpha-ketoglutarate; NADH = nicotinamide adenine dinucleotide; FADH = flavin adenine dinucleotide.

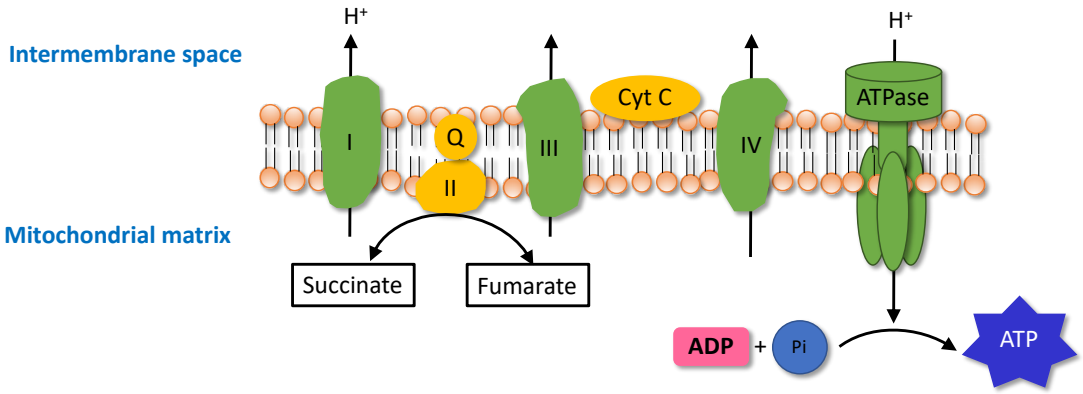


Figure 1.4. Schematic of the electron transport chain (ETC). The TCA cycle interfaces with the ETC at succinate dehydrogenase (complex II). Numerals indicate respiratory complex number. Arrows between mitochondrial matrix and intermembrane space indicate direction of proton (H^+) flow. Q = coenzyme Q10; Cyt C = cytochrome C; Pi = inorganic phosphate; ADP = adenosine diphosphate; ATP = adenosine triphosphate.

Although canonically a feature of the mitochondria, elements of the TCA cycle also occur in the cytoplasm [122] and the nucleus [123] (**Figure 1.5**). Isocitrate dehydrogenase and malate dehydrogenase both have cytoplasmic isozymes (IDH1 and MDH1, respectively), and a single form of fumarate hydratase (FH) can be found in both the cytoplasm and the mitochondria. Furthermore, during development, several mitochondrial TCA cycle enzymes including PDH, PC, and isocitrate dehydrogenase 3 (IDH3) can translocate to the nucleus, where they ultimately generate αKG [123] although to date this has only been demonstrated during a time-limited phase of zygotic genome activation.

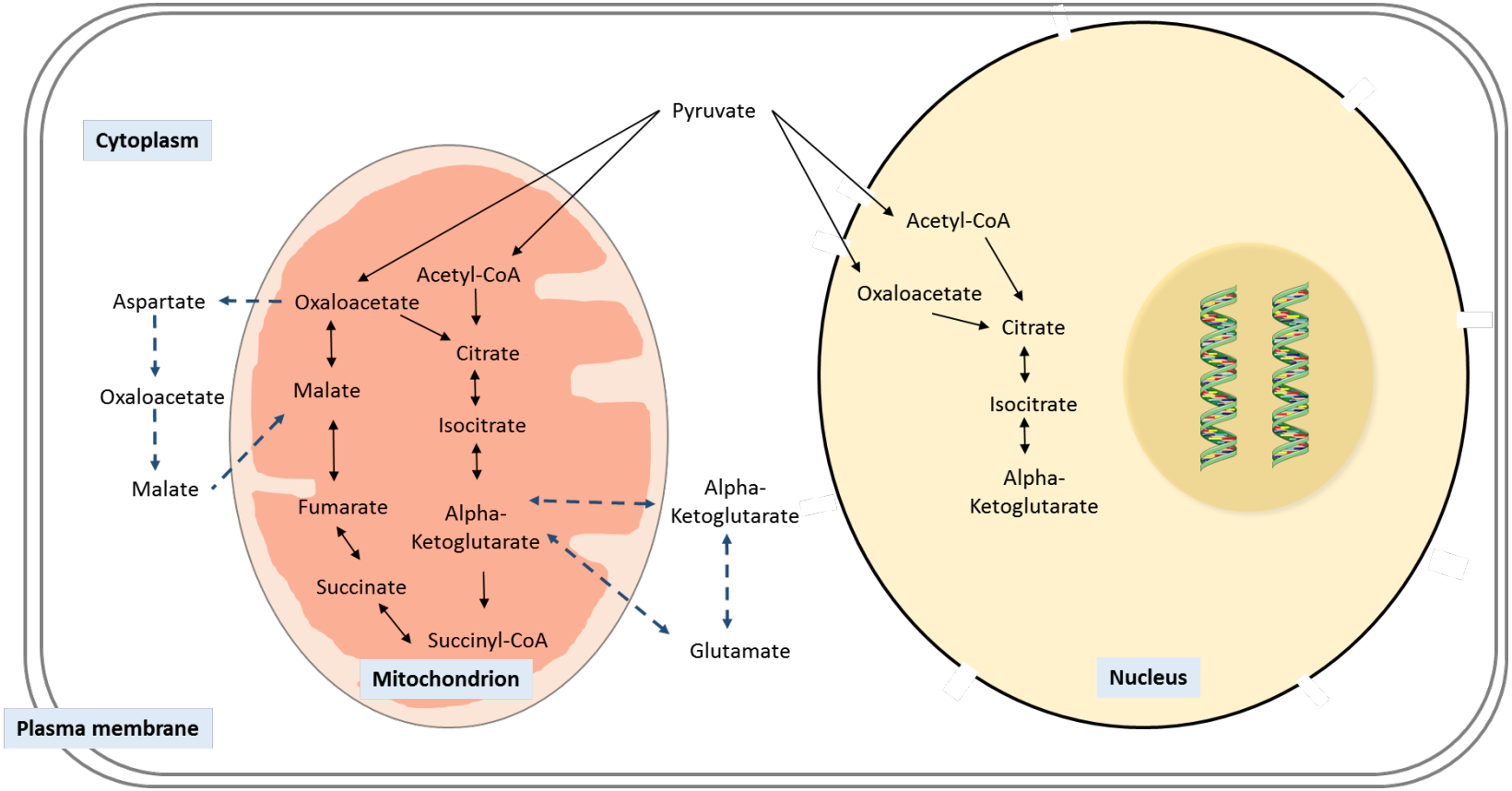


Figure 1.5. Multiple compartments of the TCA cycle. The TCA cycle occurs predominantly in the mitochondrial matrix, although some reactions can occur beyond the mitochondria, including in the cytoplasm and nucleus. Solid arrows represent the reaction direction (enzymes omitted for simplicity), while dashed arrows represent the transport of metabolites into the cytoplasm (Sinton *et al*, 2019) [124].

1.3.3 Plasticity of the TCA cycle

Whilst the standard model of the TCA cycle describes a unidirectional process, the landscape is more complex than this. Instead of OAA condensing with acetyl-CoA to form citrate, it is possible for it to instead be converted to malate, then fumarate, and back to malate and OAA, in a process termed 'backflux' [125]. Truncation of the TCA cycle is also possible in certain disease states, including various cancers and hypoglycaemia, whereby specific steps are no longer functional and are, therefore, bypassed. During hypoglycaemia, for example, it is possible for glutamate to become the primary source of carbon, allowing OAA to be generated from αKG instead of citrate [126]. Loss of FH function, as found in some cancers, results in the accumulation of fumarate, and a corresponding decrease in NADH generation. However, despite these metabolic alterations these cells are viable, and resist hypoxia, perhaps through stabilisation of hypoxia inducible factors [127], which will be discussed later in this thesis. The plasticity of the TCA cycle serves to illustrate that it can undergo complex alterations to overcome pathological states and promote cell survival.

There are strong indications that TCA cycle activity increases in response to NAFLD, in both mouse and humans, with both demonstrating increased cycle flux and anaplerotic influx of αKG *via* glutamate. Evidence for this is compelling, but the technical difficulties of obtaining human tissue samples mean that measurements of hepatic TCA cycle flux and anaplerosis in humans have been made in peripheral blood, and thus the results may be confounded by the release of metabolites from other tissues [34]. Studies of the liver of mice fed a high fat diet demonstrate increased TCA cycle flux associated with NAFLD, supporting findings in humans [110] although the use of whole tissue to examine the effects of NAFLD on metabolic function makes it difficult to determine which cells within a tissue are exhibiting altered TCA cycle activity. Supporting this, though, are nuclear magnetic resonance (NMR) studies, which also indicate increased hepatic TCA cycle flux and anaplerosis in humans [34]. A recently developed *in vitro* model, which uses embryonic stem (ES) cell-derived hepatocytes to model NAFLD, supports the assertion that intracellular lipid

accumulation associates with dysregulation of the TCA cycle and altered abundance of intermediates [128].

1.3.4 The TCA cycle and NAFLD

Satapati et al. (2012) observed increased tricarboxylic acid (TCA) cycle activity and anaplerosis in the livers of mice with high fat diet-induced hepatic steatosis [109]. They also demonstrate that increased levels of hepatic TGs induced oxidative metabolism, with a proportional increase in oxidative stress [129]. Indirect studies in humans indicate a similar effect on TCA cycle flux [34]. As TCA cycle flux generates NADH and $FADH_2$, which transfer electrons to the ETC, alterations in TCA cycling may have a direct impact on respiration and oxidative stress. Further studies in humans demonstrate that NAFLD and NASH are associated with increased and decreased mitochondrial maximal respiration, respectively, indicating that there is a transition in mitochondrial function between these two disease states [130]. The reasons for this are unclear; such differences in maximal respiration between NAFLD and NASH could reflect impairment of the ETC and/or changes in baseline mitochondrial respiration. Furthermore, mitochondria isolated from whole tissue come from multiple cell types, which may differ in proportion in disease states and this may confound the interpretation of any results. In NASH, the decrease in maximal respiration which occurs despite an increase in mitochondrial mass may be a consequence of mitochondrial leakage resulting from decreased outer membrane integrity, a phenotype associated with oxidative stress and apoptosis [130,131].

1.4 The TCA cycle and the epigenome

1.4.1 The αKG-dependent dioxygenases

In addition to its central role in energy metabolism, the TCA cycle exerts control over the catalytic activity of the αKG-dependent dioxygenase family which comprises approximately 70 members [132]. The αKG-dependent dioxygenases are $Fe(II)/O_2$-dependent enzymes, which catalyse the hydroxylation of a number of substrates,

including DNA, histones, and proteins. The family of enzymes include prolyl-4-hydroxylases (PHDs), Jumonji domain-containing histone demethylases (JHDMs), and ten-eleven translocation (TET) dioxygenases [133]. All of the enzymes within this family have two binding domains in their β-sheet core, which are highly specific for αKG - binding of αKG acts as a positive regulator of dioxygenase activity [134,135]. Conversely, these two binding domains can be competitively bound by succinate and fumarate, with subsequent inhibition of dioxygenase activity [136].

The functions of the members of the dioxygenase family are numerous, but include epigenetic regulation, through DNA methylation, histone methylation and acetylation, as well as oxygen sensing through the prolyl-4-hydroxylases (PHDs) [137–139]. It is tempting to speculate, therefore, that the αKG-dependent dioxygenases form a nexus, integrating environmental signals and disseminating them to downstream targets, altering transcription to allow cells to maintain homeostasis. Taken together, it is not surprising that there is evidence indicating that JHDM, PHD, and TET enzyme activity may be related to the pathogenesis of NAFLD, although it is less clear whether these are a cause or consequence of the disease.

1.4.2 Prolyl-4-hydroxylases (PHDs), HIF-1α and NAFLD

PHDs catalyse the most prevalent post-translational modification (PTM) in humans, hydroxylating proline residues to form (2S,4R)-4-hydroxyproline. Whilst PHDs have many substrates, including Argonaut 2 [140], collagen [141], and elastin [142], one of the most well characterised is the hypoxia inducible factor (HIF) transcription factor [143]. HIF-1 is comprised of two constitutively expressed sub-units: HIF-1α and HIF-1β [144]. Under normal oxygen conditions (normoxia), PHDs hydroxylate two proline residues (Pro402 and Pro 564) within the oxygen-dependent degradation domain (ODDD) of HIF-1α [144]. The hydroxylation of these proline residues recruits ubiquitin E3 ligase and leads to the polyubiquitination of HIF-1α, marking it for degradation by the proteasome [145]. Under hypoxic conditions, low oxygen concentrations diminish the ability of PHDs to hydroxylate these proline residues, and thus HIF-1α stabilises,

allowing it to translocate to the nucleus where it can dimerise with HIF-1β [146]. This allows the HIF complex to bind to hypoxia response elements (HREs), facilitating transcription (**Figure 1.6**).

Despite their name, it is not only oxygen that regulates the HIFs. Since PHDs are also regulated by succinate, fumarate, and αKG, HIFs are able to coordinate gene networks in response to metabolic rewiring [147]. Indeed, hypoxia is associated with altered lipid homeostasis, through coordination of genes regulating lipid droplet formation, DNL, and fatty acid uptake [148]. Whilst there is a lack of data to directly link TCA cycle dysfunction in NAFLD with PHD inhibition and HIF stabilisation, there are limited data from rat models demonstrating altered HIF-1α mRNA and protein levels in steatotic whole liver [149]. A caveat to this study is that the study was performed in whole tissue, and perivenous regions of the liver express higher levels of HIF mRNA [150]. Therefore, variation in region of the liver sampled may alter results. However, in support of this link, there are obesity studies suggesting a relationship between lipid accumulation and HIF-1α stabilisation. Adipose tissue levels of both HIF-1α mRNA and protein levels are increased by obesity, and there is strong evidence that obesity provokes a hypoxia-like response in adipose tissue [151,152]. Tracer studies in mice, using a ^{13}C-glucose stable isotope, demonstrated an increase in the abundance of αKG, succinate, and fumarate in the adipose tissue of obese animals [35]. Furthermore, this study found increased levels of succinate in the liver of obese mice, suggesting that there may be an impact on the activity of PHDs.

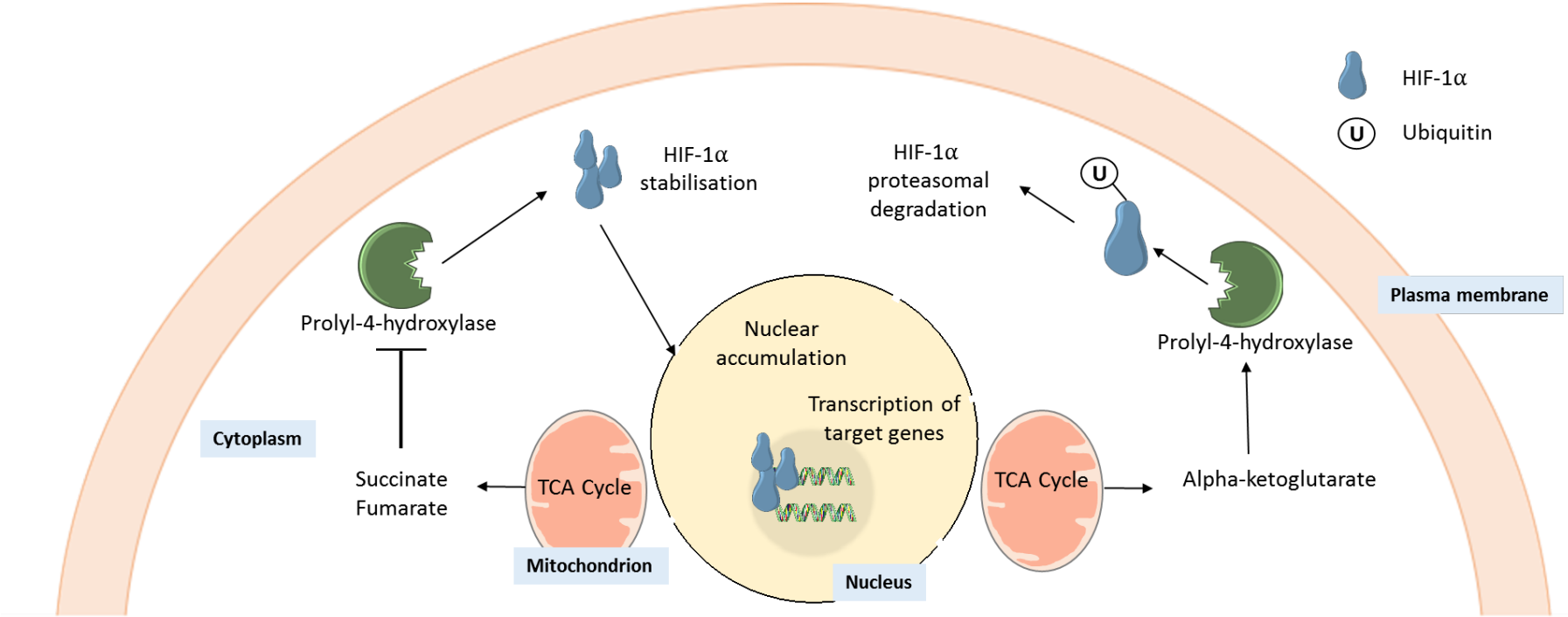


Figure 1.6. The TCA cycle and prolyl-4-hydroxylases (PHDs). TCA cycle metabolites in the cytoplasm can interact with PHDs, to either promote or repress activity. Repression prevents ubiquitination of HIF-1α, promoting protein stabilisation, followed by nuclear accumulation and transcription of target genes (Sinton *et al*, 2019) [124].

There are conflicting reports on the protective or harmful role of HIF-1α in the pathogenesis of fatty liver [149,153]. Using a rat model of liver steatosis, Carabelli *et al* describe increased expression of HIF-1α mRNA and suggest that this is responsible for an observed increase in mitochondrial biogenesis. They also speculate that this may form a protective mechanism, whereby hepatocytes are able to increase rates of β-oxidation to process excess intracellular lipids. In contrast, Mesarwi *et al* observed that HIF-1α mediates the development of hepatic fibrosis in a mouse model of NAFLD, suggesting a pathogenic role. The discrepancies between these studies may be explained by the use of different animal models. It may also be a result of NAFLD comprising a spectrum of pathology, whereby some molecular mechanisms may be protective during the early stage of disease, but eventually become chronic and harmful. However, there is a lack of data to suggest whether HIF-1α is important for the development of NAFLD in humans, and the impact of TCA cycle dysfunction on PHD catalytic activity.

1.5 Epigenetic modifications

The term epigenetics is commonly defined as the study of heritable phenotypic changes that are not related to alterations in DNA sequence, and is used to describe the molecular mechanisms that regulate transcription [154–156]. This can include histone modifications, DNA methylation, and non-coding RNAs. These modifications are able to persist through multiple cell divisions, but whether they are heritable remains a topic of debate [157,158]. It is also possible for epigenetic modifications to drive cell- or tissue-specific transcriptional programs [159], and to be dynamically influenced by diseases, such as cancer [160,161], and environmental factors [124].

1.5.1 DNA packaging and chromatin accessibility

The nucleus of a eukaryotic cell is 2-10 μm in diameter and, therefore, the several metres of DNA contained within must be packaged efficiently, whilst remaining accessible to the transcriptional machinery [162]. To facilitate both of these requirements, eukaryotic cells employ histone proteins. Approximately 147 base pairs

of DNA wrap around each histone octamer, forming nucleosome complexes, which then condense to form chromatin [163]. Chromatin is regulated by numerous factors, including chaperones, which can facilitate the deposition or eviction of histones [164,165]. Chromatin remodellers, which use ATP, can evict histones to regulate transcription [166]. It is also possible for the histones to be modified through addition of chemical tags, which will be discussed in the following section.

1.5.2 Histone modifications

Histone proteins form cores around which DNA can condense, allowing it to be packaged efficiently into nucleosomes to form chromatin [163]. These nucleosomes are often compacted, preventing transcription of DNA to mRNA [167]. Histones are also subject to many PTMs, including acetylation, methylation, phosphorylation, sumoylation, and ubiquitylation, with acetylation and methylation being the most well described [168]. Histone lysine methylation can lead to condensation of chromatin into a tightly compacted heterochromatin state, or can relax the chromatin into a euchromatic state, depending on the site of deposition [169]. Histone 3 lysine 27 trimethylation (H3K27me3) [170], histone 3 lysine 9 di- and tri-methylation (H3K9me2/3) [171,172] and histone 4 lysine 20 trimethylation (H4K20me3) are associated with the heterochromatic state and transcriptional silencing [173,174]. Conversely, histone 3 lysine 4 trimethylation (H3K4me3) and histone 3 lysine 36 trimethylation (H3K36me3) are associated with a euchromatic state and transcriptional activation [175–177]. In addition to compacting or relaxing chromatin, there is evidence that histone modifications can act as a dock for effector proteins, the binding of which can elicit a transcriptional response [178,179]. However, this may oppose the effect that the histone modification has on chromatin structure [180].

1.5.3 Jumonji domain-containing histone demethylase (JHDMs)

The JHDMs are αKG-dependent dioxygenases, which catalyse the demethylation of lysine residues of histone 3 (H3) and histone 4 (H4) [181]. In general, methylation of H3 and H4 tails is associated with relaxation of chromatin, and increased DNA accessibility [182]. In contrast, demethylation by the JHDMs can increase DNA histone interactions, promoting chromatin compaction and transcriptional silencing [183]. Experimental silencing of succinate dehydrogenase results in the accumulation of the TCA intermediate succinate, which, in turn, results in the inhibition of the JHDMs, altered histone methylation patterns, and transcriptional activity [184].

1.5.4 DNA methylation

DNA methylation is a highly conserved epigenetic modification, found in plants and animals, and is used by the cell to modulate gene expression [185]. It is essential for development and is associated with repression of transposons and X-chromosome inactivation [186]. DNA methylation occurs at the C5 position of cytosine, and is deposited by DNA methyltransferases 1-3 (DNMTs) at cytosine-phosphate-guanine (CpG) sites within the genome, giving rise to 5-methylcytosine (5mC) [187]. The presence of 5mC is associated with transcriptional repression, particularly in promoter regions [188]. The presence of 5mC at enhancers, which are distal gene regulatory elements, does not appear to be related to transcriptional repression, but rather appears to stabilise demethylated DNA, enabling transcription factor binding [189]. It is possible for 5mC to influence transcription by either physically preventing transcriptional proteins from binding [190] or through the binding of methyl-CpG-binding domain (MBD) proteins to 5mC. MBDs can recruit repressive proteins, such as histone deacetylases, to the methylated locus, generating heterochromatin [191]. The removal of the methyl group of 5mC can happen either passively or actively. The former is achieved through failure of the DNMTs to add methyl groups to the daughter strand of DNA during DNA replication. DNA methylation can be oxidised by the ten-eleven translocation (TET) enzymes to sequentially produce 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), 5-carboxylcytosine (5caC),

before triggering base excision repair (BER) to revert the base to unmodified cytosine (**Figure 1.7**) [137].

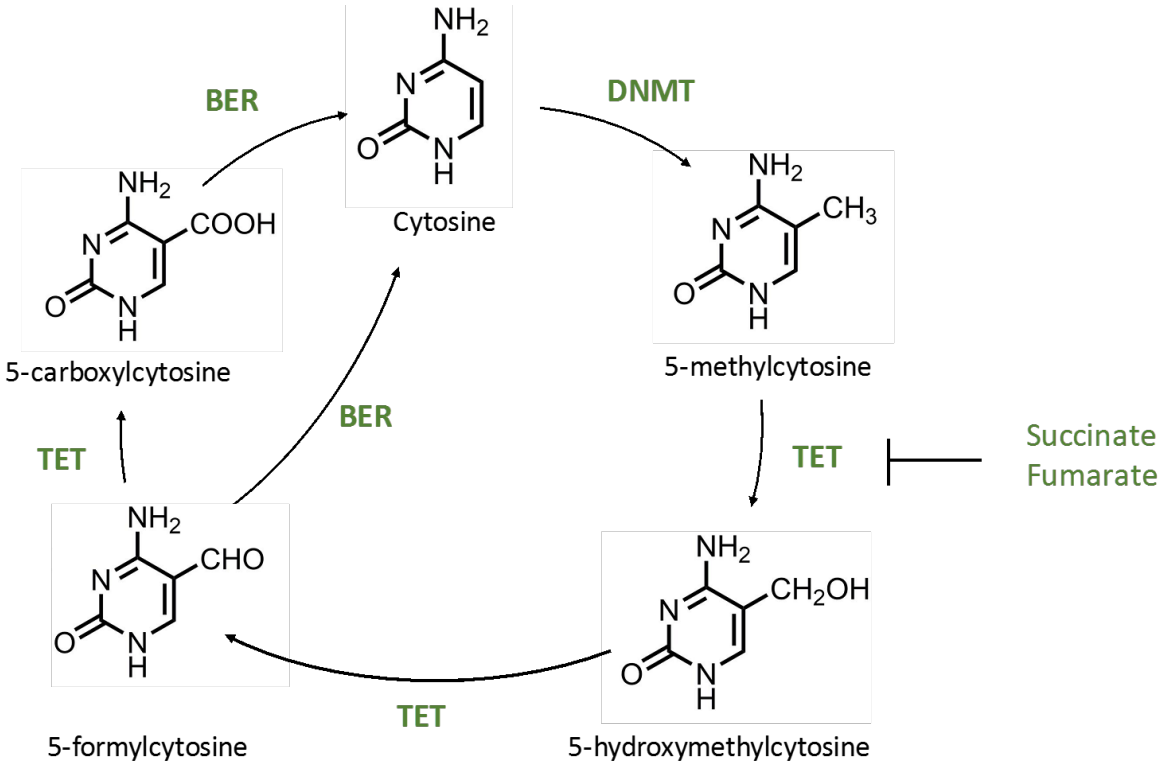


Figure 1.7. TET oxidation of 5mC through to unmodified cytosine. After the deposition of 5-methylcytosine by the DNMT enzymes it can be oxidised by the TETs to 5-hydroxymethylcytosine, 5-formylcytosine, and then 5-carboxylcytosine, before base excision repair resulting in unmodified cytosine. These iterative oxidation steps can be inhibited by the TCA cycle metabolites succinate and fumarate (Sinton *et al*, 2019) [124].

1.5.5 CpG islands

In somatic cells, >80% of CpGs are methylated, except in CpG islands and other gene regulatory regions [192]. Approximately 10% of CpG sites are located in clusters termed CpG islands (CGIs), and are most commonly found in the promoters of genes with important developmental functions [193]. In general, these CGIs are unmethylated, but where methylation does occur, it is associated with transcriptional repression and heterochromatic regions, as DNA methylation is, in general, thought

to prevent transcription factor binding [194]. During development, it is estimated that less than 30% of CpGs are methylated, in contrast to 85% in terminally differentiated cells, illustrating the tight control that is exerted to retain cell homeostasis and phenotype in differentiated cells [195].

1.5.6 DNA hydroxymethylation

Whilst initially thought to be transient, 5hmC is now accepted as a stable cytosine modification that is involved in transcriptional regulation [196,197]. It is between 10- and 100-fold more abundant than either 5fC or 5caC, and is most enriched in neurons and stem cells, whilst being more commonly depleted in cancer cells [198–200]. It is becoming increasingly evident that 5hmC is important for the maintenance of pluripotency and central nervous system development, as well as in tumorigenesis [201].

Enrichment of 5hmC within gene bodies and enhancers broadly correlates with transcriptional activation [202,203]. In contrast, 5hmC enrichment within gene transcriptional start sites (TSS) is associated with transcriptional repression and the maintenance of CpG hypomethylation [204]. 5hmC is often present in the TSS of genes where promoters are marked by the bivalent, or “poised” histone modifications H3K27me3 and H3K4me3 [205], suggesting that 5hmC is involved in modulating chromatin accessibility. There is also an enrichment of 5hmC in active enhancer regions that are labelled with H3K4me1 and the acetylated histone 3 lysine residues H3K18 (H3K18ac) and H3K27ac [206]. This demonstrates that the influence of 5hmC on transcription is context dependent, as it is able to both promote and repress mRNA expression. This complex interplay is also cell type-dependent, further complicating the role of 5hmC in transcription [207].

Altered 5hmC is also associated with different disease states or drug treatments. For example, induction of HCC in mice using the carcinogen phenobarbital dramatically

alters 5hmC at specific DNA loci. In these regions, both depletion and enrichment of 5hmC were associated with increased mRNA expression, as well as increased levels of the activating histone mark H3K4me2. In a mouse model of NAFLD, steatosis was associated with reversible changes in 5hmC at a number of functionally relevant genes, including dehydrocholesterol reductase 7 (*Dhcr7*) and 17-beta-hydroxysteroid dehydrogenase (*Hsd17b7*) [208]. This indicates that 5hmC plays a role in pathogenesis of NAFLD.

1.5.7 Ten-eleven translocation enzymes

The TET enzymes are part of the αKG- and Fe^{2+}-dependent dioxygenase family, and were first identified as homologues of the trypanosome J-binding proteins [137]. Like other members of the dioxygenase family, the TETs are allosterically regulated by metabolites of the TCA cycle, with αKG promoting activity, and succinate and fumarate inhibiting their activity [137]. Of these metabolites, αKG has the highest binding affinity, meaning that in the presence of all three metabolites it is highly likely that αKG will bind and promote TET activity [209]. To achieve this, the double-stranded β-helix at the N-terminus of the TET enzymes brings together Fe(II), αKG, and 5mC to allow oxidation of the methyl group to 5hmC, 5fC, and 5caC [137]. The contiguous cysteine-rich domain then wraps around the double-stranded β-helix to stabilise the interaction between the TET enzyme and the DNA, facilitating the oxidation reaction [137]. The enzymes are a subset of the αKG-dependent dioxygenase family, comprised of TET1, TET2, and TET3, each of which can catalyse the iterative oxidation and demethylation of 5mC to 5hmC *in vitro* and *in vivo* [137,210]. However, studies indicate that TET1, which was first identified as a fusion partner of MLL in myeloid leukaemia [211], is responsible for demethylation of promoter regions, whilst TET2 is more active at CpGs located in gene bodies [212].

The different TET isozymes show tissue-specific and temporal expression patterns [213–215]. TET1 is highly expressed in primordial germ cells and embryonic stem cells [137,216], suggesting a link to pluripotency. In contrast, TET2 is more highly

expressed in somatic cells, particularly haematopoietic cells [217]. Despite differences in expression of these two isozymes, they have overlapping roles in regulating haematopoietic homeostasis, hematopoietic stem cell maintenance, and leukemogenesis [218]. In contrast to TET1 and TET2, TET3 is most highly expressed in oocytes and zygotes, contributing to paternal global DNA demethylation [219]. Furthermore, the human TET1 and TET2 isozymes have higher affinities for 5mC than they do for 5hmC or 5fC, meaning that once 5hmC is established in genomic DNA it is less likely to be removed [220]. In contrast, TET3 can act as a 5caC reader during BER, due to the high affinity of its CXXC domain [221].

In human ES cells, TET1 may orchestrate the activity of components of the epigenetic machinery by targeting genes for binding by the polycomb recessive complex (PRC2). Though they do not appear to physically interact, TET1 appears to recruit PRC2 to specific genes [222], possibly by maintaining a hypomethylated state, as DNA methylation prevents binding by PRC2 [223]. Upon binding to hypomethylated DNA, PRC2 can di- or trimethylate H3K27, an epigenetic mark associated with transcriptional repression [170,224]. This may be relevant to the pathogenesis of NAFLD, since genome-wide association studies have shown enrichment of PRC2 pathways in humans with NASH [225]. In support of this, data from rat liver and human liver cell lines have demonstrated that downregulation of core components of PRC2 is associated with the de-repression of inflammatory genes and intracellular lipid accumulation [226]. Thus, whilst TET1 may not directly control the DNA-binding activity of PRC2 in the pathogenesis of NAFLD, these data suggest that it may facilitate it.

1.5.8 Post-translational regulation of TET activity

In addition to regulation by TCA cycle metabolites, Fe(II), and oxygen, the TET enzymes are also regulated by PTMs, including phosphorylation and *O*-GlcNAcylation. The *O*-GlcNAcylation PTM is deposited by *O*-linked N-acetylglucosamine transferase (OGT), which attaches an N-acetyl glucosamine to

serine and threonine residues within target proteins [227]. Hrit *et al* demonstrated that disruption of the OGT binding site in TET1 reduced genomic 5hmC levels, with a concurrent reduction in gene expression [228]. This indicates that in addition to the other factors described here, *O*-GlcNAcylation also impacts on TET catalytic activity.

The posttranslational addition of *O*-GlcNAcylation to proteins is promoted by the hexosamine biosynthetic pathway (HBP), which synthesises N-GlcNAc from glucose, linking elevated intracellular glucose to increased *O*-GlcNAcylation [229]. This generates a positive feedback loop, whereby the HBP diverts glucose away from glycolysis, and lowers αKG levels. This, in turn, promotes stabilisation of HIF-1α via PHD inhibition, and promotes increased transcription of the glucose transporter 1 (GLUT1) gene [230], allowing more glucose to enter the cell. The reduction of αKG in response to elevated *O*-GlcNAcylation could indicate that, whilst TCA cycle intermediates play a direct role in TET enzyme function, they may also play an indirect role through the modulation of PTMs.

In addition to its role in the deposition of *O*-GlcNAcylation PTMs on the TET enzymes, OGT also promotes the transition from hepatic steatosis to HCC, demonstrated through the use of both mouse models and human cell lines [231]. Whilst this effect is thought to be mediated by palmitic acid production and ER stress, other avenues have not been explored. OGT has been found to form a complex with TET1 or TET2 in the nucleus, anchoring it to histones and DNA [232], and since genomic hypomethylation is a hallmark of HCCs [233], it is tempting to speculate that this complex may be involved in its development.

1.5.9 Epigenetic changes in NAFLD

There is evidence to indicate that altered TCA cycle activity, associated with obesity and NAFLD, results in the repression of JHDM activity [234,235]. In a mouse model of hepatic steatosis, the JHDMs were found to alter H3K4 and H3K9 trimethylation

levels at peroxisome proliferator-activated receptor alpha (PPARα) and in lipid catabolism networks in response to intracellular lipid accumulation [236]. Whilst there is no direct evidence to indicate whether this also occurs in humans, these findings are highly suggestive that similar mechanisms may be at play, making the JHDMs attractive targets to study in the context of NAFLD.

Changes in hepatic DNA methylation may be associated with the development of steatosis, as demonstrated by Pogribny *et al*. Using a methionine-deficient diet (MDD) model of NAFLD, in mouse, they observed decreased expression of DNMT1, and a concurrent decrease in global levels of DNA methylation [237]. However, the authors do not describe in detail where in the genome these decreases occur, making it difficult to conclude the impact that this may have on transcriptional regulation. Furthermore, there are limitations to recapitulating hepatic steatosis with the MDD model, as shown by suppressed expression of very low density lipoprotein carboxylesterase mRNA [238]. In contrast, Pirola *et al* observed that humans with NASH have increased DNMT1 mRNA, and a concurrent increase in NADH dehydrogenase 6 methylation [239]. Whilst interesting, the findings are correlative, and do not describe a mechanism by which altered methylation occurs. In addition, the authors used bisulfite treatment to measure methylation levels, making it impossible to determine if they were observing changes in 5mC or its oxidised counterpart 5hmC.

Evidence of altered nuclear DNA methylation in the liver in NAFLD is supported by a number of studies in humans. Differential methylation has been reported in liver biopsy specimens from patients with NAFLD [240,241], with some studies showing an inverse correlation between gene expression and DNA methylation [242]. Specific differentially methylated regions have been identified which may associate with NAFLD progression [243]. Notably, some of the changes in DNA methylation may be reversible with weight loss, for example following bariatric surgery [242,243]. There is also evidence from bariatric surgery patients suggesting that there is a correlation between altered DNA methylation and changes in hepatic insulin signalling [244]. A

recent study focused on understanding the link between NAFLD-associated liver fibrosis and cirrhosis and DNA methylation performed bisulfite sequencing in liver biopsies from female patients undergoing bariatric surgery, who had no other liver-related pathologies [245]. The study identified differential CpG island methylation at a number of different genes and gene ontology analysis showed enrichment for a number of pathways related to NAFLD, including reactive oxygen species production and T2D signalling [245]. Whilst compelling, the findings in each of these studies are correlative and do not describe any mechanism(s) by which altered DNA methylation might occur. In addition, the studies again used bisulfite treatment to analyse DNA methylation, making it impossible to determine if they were observing changes in 5mC or 5hmC.

There is evidence to suggest that 5hmC levels are altered in NAFLD. Patient studies describe a positive association between global hepatic 5hmC levels and the expression of peroxisome proliferator-activated receptor γ (PPARγ) coactivator 1α (PPARGC1A), a gene at which DNA methylation has previously been associated with NAFLD and IR [239][246]. The lack of any change in global 5hmC could reflect the use of whole tissue, which may dilute the 5hmC signal through analysis of multiple cell types or alternatively may indicate that alterations in 5hmC in response to NAFLD occur only at specific genes. Indeed, using an in vitro model of NAFLD, it was recently shown that changes in 5hmC levels occur specifically at induced genes associated with lipid synthesis and transport [128]. Further, the study found altered expression of several genes encoding TCA cycle enzymes, including succinate dehydrogenase, isocitrate dehydrogenase, and an accumulation of the TCA cycle metabolite oxalosuccinate, suggesting altered cycle activity. The importance of the TET enzymes and 5hmC in liver disease is further supported by studies showing that aberrant TET1 activity in HCC is associated with loss of 5hmC and a gain of 5mC, with subsequent transcriptional repression, suggesting that the TETs may normally prevent DNMT binding in these regions in the healthy liver [247].

1.6 Modelling NAFLD *in vitro*

Many studies have modelled NAFLD using either rodents or *in vitro* culture systems, in an attempt to understand the pathophysiology underpinning this spectrum of diseases. However, the multi-factorial nature of NAFLD makes it challenging to accurately model using these systems. For example, mouse models of NAFLD often induce weight gain and hepatic steatosis through feeding of a high saturated fat diet [248,249]. When combined with sucrose in the diet, this can recapitulate some elements of NASH [250]. Further to this approach, is the MDD diet, described in section 1.5.9, which recapitulates the inflammatory response associated with NASH, but not the other elements of the metabolic syndrome [238,251]. Beyond dietary interventions is the use of genetic manipulation to increase appetite, such as in Ob/Ob or Db/Db mice [252]. Alternative genetic approaches have been used to reduce VLDL export ($MTT^{-/-}$) or fatty acid oxidation ($Ppara^{-/-}$) [252]. A challenge of using rodent models to study metabolism of distinct cell populations is that tissue must first be disaggregated and then sorted, which induces oxidative stress, potentially confounding analysis [253].

In vitro approaches commonly utilise immortalised HCC cell lines, such as HepG2s. However, malignant transformations, in addition to immortalisation, are known to induce profound changes in metabolic phenotype [254,255], limiting their utility as models of NAFLD. Furthermore, culturing of HCC cell lines is associated with loss of 5hmC, making them unsuitable for studying the interplay between mitochondrial metabolism and modulation of the epigenome [247,256–259]. As an alternative, primary mouse and human hepatocytes have been used to model NAFLD [260]. However, these have a limited lifespan and cannot be passaged or expanded *in vitro*. Furthermore, as with rodent studies, analysis of primary cells are confounded by the cell sorting process [253].

In recent years, limitations of these *in vivo* and *in vitro* systems has led to an increased interest in developing human pluripotent stem cell (hPSC)-derived hepatocyte models

of NAFLD, using either adult somatic cells (induced pluripotent stem cells (iPSCs)) or embryonically-derived stem cells (ESCs) [261–264]. Whilst iPSCs are attractive, due to their potential use in personalised medicine, they are challenging to derive from patients and require phenotyping before use [265]. Furthermore, some iPSCs exhibit incomplete programming and proliferative capacity [266]. The use of ESC-derived hPSCs to model NAFLD is more straightforward. Whilst these cells do require regular phenotyping, to assess pluripotency and genome integrity, they have much greater proliferative capacity than iPSCs and can be used for high-throughput studies [267]. Recently, an ESC-derived hPSC model was leveraged to model NAFLD *in vitro*, whereby human H9 cells were differentiated to hepatocyte-like cells (HLCs) and exposed to a cocktail of high energy substrates: lactate, pyruvate and octanoic acid (LPO), inducing lipid accumulation [128]. Previous studies used a similar cocktail, with the addition of ammonia, to induce lipid accumulation in an HCC cell line [42]. However, the mechanisms underpinning lipid accumulation in response to LPO exposure remain unclear. Exposure of HLCs to LPO induced transcriptional changes similar to those observed in human disease, including in genes related to the TCA cycle and lipid metabolism. Furthermore, untreated HLCs displayed 5hmC levels similar to those observed in adult human hepatocytes, which were modulated in response to LPO. These findings demonstrate that this model recapitulates the metabolic and epigenetic features of NAFLD, enabling the study of hepatocyte response to lipid accumulation in a single cell population. This model was refined and used for the studies described in this thesis.

1.7 Summary, hypotheses and thesis objectives

The study of epigenetic dysregulation in disease is fraught with difficulties and limitations, not least whether changes in epigenetic marks are a cause or consequence of the disease process itself [268]. In addition to this, many studies in the NAFLD field have reported changes in genome methylation but have arrived at their conclusions based on bisulfite sequencing data [239,242]. Since bisulfite sequencing cannot distinguish between 5mC and 5hmC [269], it is impossible to determine precisely which changes are occurring during the pathogenesis of NAFLD

in much of the current literature. Further to these difficulties are the inherent limitations of the current models of NAFLD, which make it difficult to study the interplay between the TCA cycle and epigenetic modifications. The aim of my thesis was to refine the hPSC-based model of NAFLD described above and to study the interplay between TCA cycle dysfunction and epigenetic regulation of gene expression in the pathogenesis of the disease.

1.8 Hypotheses and aims

1.8.1 Hypotheses

1. LPO exposed HLCs are transcriptionally similar to steatotic hepatocytes obtained from patients with NAFLD.
2. LPO increases TCA cycle activity, leading to increased levels of the metabolites αKG, succinate and fumarate.
3. Increased αKG levels will increase TET enzyme activity, modulating levels of 5hmC throughout the genome.

1.8.2 Aims

1. To examine the transcriptional response of LPO-treated HLCs in response to model refinement using an RNA-sequencing approach.
2. To perform high-resolution analyses of mitochondrial respiratory changes associated with intracellular lipid accumulation.
3. To determine whether LPO-induced changes in mitochondrial TCA cycle activity impact on TET enzyme activity.

Chapter 2: Materials and Methods

2.1 Buffers and solutions

Reagents were purchased from Sigma Aldrich, except where stated.

10x Western Blot Running Buffer: Tris Base 121 g, sodium dodecyl sulphate 10 g, HEPES 239 g were dissolved in 1 L deionised water and used at 1 x concentration.

Western Blot Transfer Buffer: Tris Base 3 g, Glycine 14.4g , 100% methanol 100 mL were dissolved in 1 L deionised water.

RIPA lysis buffer with protease inhibitors: 150 mM sodium chloride, 1x Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, and 50 mM Tris (pH 8.0) were dissolved in deionised water. 10ml aliquots were stored at -20 °C. cOmplete™ Protease Inhibitor Cocktail tablets were added (1/10 mL buffer; Roche, 11697498001).

10x TBE buffer: Tris Base 108 g, boric acid 55 g and 40 mL of 0.5 M EDTA solution (pH 8.0) were dissolved in 1 L deionised water. This was used at 1 x concentration.

1M Na-Phosphate (mono and dibasic): 1 M solutions of mono and dibasic sodium phosphate salt were made up separately in distilled water, combined in a 39:61 ratio and adjusted to pH 7.0 with 4 M NaOH.

DNA Immunoprecipitation Buffer: 100 mM Na-Phosphate pH 7.0 (mono and dibasic), 1.4 M NaCl, 0.5 % Triton X-100 were made up to 10 mL in deionised water fresh and 0.22 μm sterile filtered.

Proteinase K digestion buffer: 50 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.5% SDS were prepared in 10 mL deionised water prior to each procedure and 0.22μm sterile filtered.

Cold 50% trichloroacetic acid solution: 25 g of crystalline trichloroacetic acid was dissolved in 50 mL deionised water and stored at 4 °C.

1% acetic acid solution: 10 mL of 100% acetic acid was dissolved in deionised water to a final volume of 1 L.

Sulforhodamine B 0.4% solution in 1% acetic acid: 0.4g of Sulforhodamine B powder was dissolved in 100 mL 1% acetic acid solution and stored at 4 °C.

10mM Tris solution pH 10.5: 1.21g of Tris base was dissolved in deionised water. The pH of the solution was adjusted to 10.5 with HCl water added to a final volume of 1 L before storing at 4 °C.

LPO Treatment: Sodium L Lactate (Sigma cat. number L7022) was made up in 1 mL aliquots of 1 M solution and stored at -20 °C. On the day of use, 11 mg of Sodium Pyruvate Sigma cat. number P5280) was added to a thawed 1 mL aliquot and kept on ice. Octanoic Acid (Sigma cat. number C2875) 865.2mg was added to 60 mL of deionised water and adjusted to pH 7.0-7.9 with NaOH 4 M solution with constant stirring. The solution clears at this point and is stored in aliquots at -20 °C. The required volume of each solution was then added to media 0.22 μM filtered and added to the culture dish. For isotopic tracing studies, lactate was replaced with $^{13}C_3$-lactate (CK Isotopes, CLM-1579-05). For mechanistic studies, HLCs were exposed to either 5-Aminoimidazole-4-carboxamide-1-β-D-ribofuranosyl 5′-monophosphate (AICAR; 1mM; Sigma-Aldrich, A1393-50MG), O-(Carboxymethyl)hydroxylamine hemihydrochloride (AOA; 100 μM ; Sigma-Aldrich, C13408-1G) or AICAR combined with monomethyl fumarate (50 μM; Sigma-Aldrich, 651419-1G) for the same duration as LPO.

2.2 Stem cell culture and differentiation

All cell culture reagents were GIBCO products supplied by Thermo Fisher Scientific unless stated otherwise. Corning (UK) supplied the plastic ware utilised throughout the cell culture. For high throughput imaging, specific μClear 96-well plates from Greiner Bio-one were used. Cell culture medium used in this study is described in Table 2.1.

2.2.1 Stem cell culture

H9 human embryonic stem cells (hESCs) were cultured on 5 µg/cm^2 of Laminin 521 (BioLamina) with mTser1 media (STEMCELL Technologies). Medium was changed daily. The cells were incubated at 37 °C in 5% (v/v) CO_2, 95% (v/v) air, for optimal growth.

2.2.2 Stem cell passaging

hESC were passaged when they reached 80% confluence. Medium was aspirated and cells were washed with Dulbecco's phosphate-buffered saline (DPBS) (Life Technologies) once. Gentle Cell Dissociation reagent (STEMCELL Technologies) was added to the cells. For regular passaging, cells were incubated for 5 min until the edges of the colonies appeared rounded. The Gentle Cell Dissociation regent was aspirated and fresh mTser1 medium was added. Using a cell scraper, a homogeneous suspension of cells was formed. The cell split ratio was 1:3 for regular colonies maintenance.

2.2.3 Stem cell freezing and thawing

Stem-cellbanker (Amsbio) was used as a freezing solution. At 80-90% confluence, hESCs were dissociated and collected, cells were centrifuged at 200xg for 3 minutes in a 15 ml tube. Supernatant was removed and the cell pellet was resuspended in 1 mL of freezing solution before transferring to -80 °C overnight and then storing in liquid nitrogen. Cells were thawed by warming the cryotube in a 37 °C water bath and then transferring to a 15 mL tube with 5 mL of fresh mTeSR1 medium to resuscitate the cells. The cells were then centrifuged at 200 x g for 3 min. The supernatant was removed, and the cells were resuspended in 2 mL of mTeSR1 medium supplemented with 10 µM ROCK inhibitor Y27632 and seeded onto a laminin 521 coated well.

2.2.4 Hepatocyte-like cell differentiation

All media used for this protocol are described in Table 2.1. For hepatocyte-like cell (HLC) differentiation, cells were incubated for 7-9 minutes with Gentle Cell Dissociation reagent until the edges of all cells were rounded, but the cells were still attached to the plate. Using a cell scraper, a homogeneous suspension of single cells was formed. Single cells, at 50,000 cells/cm^2, were plated on wells coated with 5 µg/cm^2 Laminin 521. When they reached 30 - 40% confluence the differentiation protocol was initiated, and mTseR media was replaced with the endoderm differentiation medium. On day 4, the second stage, hepatoblast differentiation medium was added for 5 days changing the media every two days. On day 9, the HLCs medium was added. The medium was changed every second day.

Cell Culture	Medium	Medium supplements	Supplier
Human embryonic stem cells (hESCs)	mTeSR1	10 µM ROCK inhibitor only for differentiation and resuscitation	Stem Cell Technologies
hESC-derived hepatic endoderm medium	RPMI 1640	B27 100ng/ml Activin A 50ng/ml Wnt3a 1% Penicillin Streptomycin	Thermo Fisher Scientific R & D System Peprotech Thermo Fisher Scientific
hESC-derived hepatoblast medium	Knockout DMEM	20% Knockout Serum Replacement (KSR) 0.5% Glutamax 1% Non-Essential Amino Acids 1% Penicillin Streptomycin 0.1mM β-Mercaptoethanol 1% DMSO	Thermo Fisher Scientific Sigma Aldrich
hESC-derived hepatocyte-like cell medium	HepatoZYME medium	1% Glutamax 1% Penicillin Streptomycin 10ng/ml hepatocyte growth factor (HGF) 20ng/ml oncostatin M (OSM) 10 µM hydrocortisone 21-hemisuccinate	Thermo Fisher Scientific Peprotech Sigma Aldrich

Table 2.1. Cell culture medium used for hESC maintenance and differentiation

2.3 Stem cell characterisation

2.3.1 Fluorescence activated cell sorting (FACS)

Fluorescence activated cell sorting (FACS) was performed to characterise the expression of cell surface proteins in hESCs. hESCs were washed once with 2 mL DPBS and then dissociated with 1mL TrypLE. Cells were collected as single cells. Post centrifugation, cells were resuspended in DPBS and filtered through a 0.22 µM filter. The single cell suspension was then incubated for 30 min at 4 °C with the fluorochrome conjugated antibodies (for antibody details, see Table 2). IgG controls were used for antibody specificity. After incubation, cells were then washed twice with DPBS, and spun down at 200 x g for 5 min. Cells were then resuspended in 300 µL of DPBS. The analysis was carried out by using a live gate on forward scatter and side scatter parameters. Data for 20,000-50,000 'live' events were acquired for each sample using a BD LSR Fortessa (4 laser) analyser. The data was analysed using FlowJo software.

Antibody	Fluorophore	Host species	Dilution	Supplier
SSEA-1	FITC	Mouse	1/50	Biolegend
SSEA-4	FITC	Mouse	1/50	Biolegend
TRA-1-60	PE	Mouse	1/50	Biolegend
TRA-1-81	PE	Mouse	1/50	Biolegend
Mouse IgG	FITC/PE	Rabbit	1/50	Biolegend

Table 2.2 Antibodies used for flow cytometry

2.3.2 CYP P450 activity

CYP 3A4 activity was measured on HLCs at day 18 by using the P450-Glo™ assay (Promega). CYP 3A4 (1:20) substrates were diluted in hepatocyte maturation medium and added to the HLCs, before incubating for 5 min at 37 °C in 5% (v/v) CO_2. Plain medium with substrate was incubated in parallel to subtract background. After incubation, 50 µL of supernatant was collected and mixed with 50 µL of luciferin detection reagent in a white 96-well plate at room temperature in the dark for 20 min. Luminescence was measured using a multiplex plate reader (GloMax Explorer, Promega) and normalised per mg of protein (RLU/ml/mg).

2.3.3 Protein content quantification

The Pierce bicinchoninic acid (BCA) assay kit (Thermo Fisher Scientific) was used to quantify the protein content in HLCs. Protein supernatant was obtained by adding 50 µL of RIPA buffer (Thermo Fisher Scientific) per well of a 96-well plate. 10 µL per samples were analysed in duplicates in a 96-well plate. A standard curve was generated using bovine serum albumin standards at concentrations ranging from 0-2000 µg/mL. Reagents A and B from the kit were mixed at a 1:50 ratio for a final volume of 200 µL per well. Samples were incubated at room temperature in the dark for 30 min and absorbance was read at 562 nm using a multiplex plate reader (GloMax Explorer, Promega). Concentrations were then calculated using the generated standard curve.

2.3.4 High content analysis microscopy

Cells were stained with a cell paint assay, adapted from Lyall et al and Bray et al [17,19]. Cells were fixed with 50 µL/well 4% (wt/vol) paraformaldehyde (PFA; Electron Microscopy Sciences, 15710-S) for 15 min at room temperature. For permeabilisation, cells were incubated in 0.1% Triton X-100 (Sigma-Aldrich, T8787) in PBS for 15 min at room temperature. For lipid droplet analysis, cells were then stained with a combination of NucBlue Live ReadyProbes® Reagent (2 drops/mL) (Molecular Probes, R37605), HCS CellMask™ Red (2 µL/10 mL) (Invitrogen, H32712), and BODIPY™ 493/503 (1:1000) (Life Sciences, D3922), as per the manufacturer's instructions. For quantification of mitochondrial fluorescence, cells were stained with MitoTracker™ Deep Red FM (1/4000; Invitrogen, M22426) prior to fixation with 50 µL/well 4% (wt/vol) PFA (Electron Microscopy Sciences, 15710-S) for 15 min at room temperature. Cells were then stained with a combination of NucBlue Live ReadyProbes® Reagent (2 drops/mL), and Alexa Fluor™ 555 Phalloidin (Invitrogen, A34055), as per the manufacturer's instructions. Following staining, images were acquired using an Operetta High Content Analysis microscope (Perkin Elmer, Buckinghamshire, UK). Lipid droplet morphology was analysed as previously described [17,20]. To do this, several random fields in each well were imaged before performing cell segmentation. Cell segmentation enables identification of features such as organelles and was performed by identifying nuclei followed by the cytoplasm.

Following this, quality control was performed and cells that were not segmented correctly were excluded, based on area, roundness and intensity. Finally, intensity, morphology and texture of cell features were quantified for each channel, for either lipid droplets or mitochondria.

2.4 Molecular biology

2.4.1 RNA Purification

Plate growing surfaces were washed twice with PBS, before incubating in QIAzol (#79306, QIAGEN) for 5 minutes at room temperature. RNA was then purified using the Monarch® Total RNA Miniprep Kit (New England BioLabs, T2010) as per the manufacturer's instructions. A DNase treatment step was performed, using the TURBO DNA-free™ Kit (#AM1907, Ambion), as per the manufacturer's instructions. Samples were snap frozen on dry ice and stored at -80°C until required.

2.4.2 RNA quantification and quality control

RNA was quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific). For RNA-sequencing (RNA-seq), sample integrity was measured using a 2100 Bioanalyzer (Agilent), to measure the RNA Integrity Number (RIN) and 28S/18S ratio. The RIN of all samples was >9.0, and the 28S/18S ratio of all samples was >1.0, meeting the requirements for sample submission for RNA-seq. For non-RNA-seq downstream analyses, A260/280 and A260/230 ratios were measured on a NanoDrop (Thermo Fisher Scientific). RNA integrity was assessed on a 1% agarose gel, stained with Gel Red (Cambridge Bioscience, Cambridge, UK) 0.005% final concentration, and assessing the presence of bands representing 28S and 18S ribosomal subunits.

2.4.3 Reverse transcription (RT)

For qPCR analysis, 1μg of RNA was added to nuclease-free H2O (NFH2O) to a volume of 10μL, which was then reverse transcribed to complementary DNA (cDNA) using the High Capacity cDNA Reverse Transcriptase Kit (#4368814, Applied

Biosystems) in accordance with the manufacturer's guidelines. 10 μL of the 2 x master mix described in Table 2.3 was added to the 10 μL of diluted RNA. This was then reverse transcribed on a Techne Primer Thermal Cycler (Techne, UK) using the cycling conditions described in Table 2.4. A negative no transcriptase control was included.

Component	Volume/Sample (μL)
10 x RT Buffer	2.0
dNTP Mix, 100mM	0.8
10 x RT Random Primers	2.0
MultiScribe™ Reverse Transcriptase	1.0
RNase Inhibitor	1.0
NFH_2O	3.2

Table 2.3. Master mix recipe for reverse transcription of RNA to cDNA

Step	1	2	3	4
Temperature (°C)	25	37	85	4
Time (min)	10	120	5	Hold

Table 2.4. Cycling conditions for reverse transcription

2.4.4 Reverse transcriction-quantitative polymerase chain reaction (RT-qPCR)

2 µL of the reverse-transcribed cDNA was used in the reaction mixture shown in Table 2.5. A list of all primer sequences used for RT-qPCR analysis of gene expression is shown in Table 2.6. A water and non-reverse transcriptase control were routinely used. All assays were carried out in triplicate. Once the mixture was pipetted into a 384-well PCR plate (Thermo Fisher, West Sussex, UK), the plate was centrifuged for 2 min at 563 x g to ensure all liquid was mixed at the bottom of the well. The plate

was inserted into the LightCycler® 480 Instrument (Roche, West Sussex, UK), and run on the program shown in Table 2.7.

The LC480 Software (Roche, West Sussex, UK) measured fluorescence throughout the experiment and plotted fluorescence against PCR cycle number. The crossing point (Cp) was calculated as the maximum point of the second derivative of the amplification curve. Triplicates were considered acceptable if the standard deviation of the Cp was less than 0.3 cycles. Negative controls were considered acceptable if the Cp of any signal was greater than 10 cycles higher than that of the lowest standard. Each assay was considered acceptable if the reaction efficiency was between 1.8 and 2.1. Relative levels of mRNA of unknown samples were interpolated from the standard curve and expressed relative to the abundance of a reference gene. Student's t-tests were carried out to ensure transcript quantities of reference genes were not significantly different between groups.

Component	**Volume/Sample (μL)**
PerfeCTa® FastMix® II (#733-2110, Quanta Biosciences Inc.™)	5.0
Forward Primer	0.4
Reverse Primer	0.4
NFH_2O	2.1
UPL Probe	0.1
cDNA	2.0

Table 2.5. qPCR Master Mix Recipe

Gene	Forward sequence	Reverse sequence	Probe
TBP	GAACATCATGGATCAGAACAACA	ATAGGGATTCCGGGAGTCAT	87
NANOG	ATGCCTCACACGGAGACTGT	CAGGGCTGTCCTGAATAAGC	69
ALB	GTGAGGTTGCTCATCGGTTT	GAGCAAAGGCAATCAACACC	7
HNF4A	AGCAACGGACAGATGTGTGA	TCAGACCCTGAGCCACCT	27
PCK2	CGAAAGCTCCCCAAGTACAA	GCTCTCTACTCGTGCCACATC	20
G6PD	AACAGAGTGAGCCCTTCTTCA	GGAGGCTGCATCATCGTACT	5
PLIN2	TCAGCTCCATTCTACTGTTCACC	CCTGAATTTTCTGATTGGCACT	72
HK1	GACCAAGTTTCTCTCTCAGATCG	CCTAGCTGCTGGAGGATAGC	1
HIF1A	GATAGCAAGACTTTCCTCAGTCG	TGGCTCATATCCCATCAATTC	64
PLIN4	AGTTCCAAGCCAGGGACAC	TGCTGGGCCTTTTCAATC	1
PLIN5	TACAGTGCAGCCAAGGACAG	CGCACACGCAGTTCTCAG	3

Table 2.6 Primer pairs and probes used for qPCR. All sequences are written 5'-3'

Step	Temperature (°C)	Time	No. Cycles
Preincubate	95	5 minutes	1
Program	95 60	10 seconds 30 seconds	50
Cool	40	30 seconds	1

Table 2.7. qPCR cycling conditions for UPL assay

2.4.5 RNA sequencing

Where RNA samples had a RIN of >8.0, they were deemed suitable for RNA sequencing. Library preparation and RNA-seq were performed by the Beijing Genomics Institute (BGI, Beijing, China). RNA samples were first depleted of ribosomal RNA (rRNA) using the Ribo-Zero Gold rRNA Removal Kit (#MRZG12324, Illumina). Subsequent library preparation was performed using the TruSeq® Stranded mRNA Library Preparation Kit (#RS-122-2101, Illumina). Samples were then sequenced using the HiSeq 4000 PE100 platform (Illumina), to a depth of >60 million reads, to provide 100bp paired end reads.

2.4.6 3'RNA-sequencing

Sequencing was performed at the Wellcome Trust Clinical Research Facility (University of Edinburgh). Library preparation was done using the QuantSeq 3' mRNA-Seq Library Prep Kit (Lexogen, Austria). The template was prepared using the Ion PI Hi-Q OT2 200 kit (Thermo Fisher Scientific, UK). Sequencing was performed using the Ion PI Chip Kit v3 (Thermo Fisher Scientific, UK).

2.5 Studying mitochondrial respiration

2.5.1 Protein Extraction

Adherent HLCs were washed once with ice-cold PBS, before incubating in ice-cold RIPA Lysis and Extraction Buffer (Thermo Scientific, 89900) supplemented with cOmplete™ Protease Inhibitor Cocktail tablets (1/10 mL buffer; Roche, 11697498001). The suspended HLCs were placed on ice for 30 minutes, vortexing every 3 minutes, before centrifuging for 20 min at 4 °C, 12,000 rpm. The supernatant was collected and stored at -80 °C until needed.

2.5.2 Western blot analysis

Protein quantification was performed using the Qubit™ Protein Assay Kit (Invitrogen, Q33211), as per the manufacturer's instructions. Protein concentration was measured

using a Qubit™ Fluorometer (Invitrogen, Massachusetts, USA). Equal concentrations (50 µg) of HLC protein extract in 4 x Sample Loading Buffer (Li-Cor, 928-40004) were loaded onto NuPAGE™ 4-12% Bis-Tris Protein Gels (Invitrogen, NP0326BOX). Following resolution, protein was transferred to a methanol-activated polyvinylidene difluoride (PVDF) membrane. Protein transfer was measured using Revert 700 Total Protein Stain Kit (Li-Cor, 926-11010) as per the manufacturer's instructions. Membranes were then blocked with Tris-buffered saline containing Tween 20 (TBST) and 5% skimmed milk powder, before incubating with either Pyruvate Dehydrogenase (C54G1) Rabbit mAb (Cell Signaling Technology, 3205) or SDHA (D6J9M) XP® Rabbit mAb (Cell Signaling Technology, 11998), both a 1:1000 dilution. The membranes were washed in TBST before incubating with the secondary antibody, IRDye® 680RD Donkey anti-Mouse IgG (Li-Cor, 926-68072) at a 1:10,000 dilution, for 1 h at room temperature, in the dark, with shaking. Blots were visualised on a Li-Cor Odyssey® CLx (Li-Cor, Nebraska, USA), and bands normalised to the Revert 700 Total Protein Stain, as per the manufacturer's instructions.

2.5.3 Citrate synthase activity assay

Citrate synthase activity was assessed as a readout of mitochondrial integrity [270,271]. Mitochondria were isolated using the Mitochondria Isolation Kit for Cultured Cells (Thermo Scientific, 89874), as per the manufacturer's instructions, selecting option A for isolation. Citrate synthase activity, a marker of mitochondrial integrity, was then measured using the Citrate Synthase Activity Colorimetric Assay Kit (BioVision, K318), as per the manufacturer's instructions.

2.5.4 Cell mitochondrial stress test assay

HLCs were differentiated directly onto Seahorse XF24 V28 cell culture plates (Agilent, 100882-004). Control and LPO groups were assigned to each well randomly, to ensure positioning did not impact mitochondrial respiration. The oxygen consumption rate (OCR) of LPO-exposed HLCs was measured using the Agilent Seahorse XF Cell Mito Stress Test Kit (Agilent, 103015-100) on a Seahorse XF Analyser (Agilent,

California, USA). Data were collected from two separate plates, representing two technical replicates. Analyses were performed under basal conditions and following treatment with oligomycin A (4 μM; an ATPase inhibitor), carbonyl-cyanide-4-(trifluoromethoxy) phenylhydrazone (FCCP; an ETC uncoupler), and combined rotenone (200 nM) and antimycin A (1 μM) (inhibitors of complex I and III, respectively). Two concentrations of FCCP (0.5 μM and 1.0 μM) were used for optimisation. Since replicates within each group responded similarly to each other, results were combined. OCR was normalised to total protein for each well, using the sulforhodamine B (SRB) assay.

2.5.5 Sulforhodamine B (SRB) assay

This was performed as previously described [272], but with spectrophotometric measurements read at 540 nm. After measuring oxygen consumption, 50 μL of 50% TCA solution was added to each well of the Seahorse XF24 V28 plate. Plates were then incubated at 4°C for 1 h. Wells were then washed 10 times with tap water and air dried, before adding 50 μL SRB to each well for 30 min at room temperature. Excess dye was removed by washing 4 times in 1% acetic acid solution. The plates were then air dried again. Cell-bound SRB was then dissolved using 200 μL tris solution, and the solution transferred to 96-well plates for spectrophotometry. Absorbance was then measured at 540 nm on a spectrophotometer.

2.5.6 Nuclear magnetic resonance (NMR) Spectroscopy

This protocol was previously described by Hollinshead *et al* [273]. NMR spectroscopy was performed by Christian Ludwig at the University of Birmingham (Birmingham, UK). At the conclusion of tracer experiments, cells were washed with 2 mL ice-cold 0.9% saline solution and quenched with 0.3 mL pre-chilled methanol (-20 °C). After adding an equal volume of ice-cold HPLC-grade water containing 1 μg/mL D6-glutaric acid (C/D/N Isotopes Inc), cells were collected with a cell scraper and transferred to tubes containing 0.3 mL of chloroform (-20 °C). The extracts were shaken at 1400 rpm for 20 min at 4 °C and centrifuged at 16,000 x g for 5 min at 4 °C. Then, 0.3 mL

of the upper aqueous phase was collected and evaporated in eppendorfs, under a vacuum using a Savant™ SpeedVac™ Concentrator (ThermoFisher). These samples were used either for NMR spectroscopy of for GC-MS. For NMR, dried samples were re-suspended in 60 μL of 100 mM sodium phosphate buffer (pH 7.0) containing 500 μM DSS and 2 mM Imidazole, 10% D20, pH 7.0. Samples were vortexed, sonicated (5-15 min) and centrifuged briefly, before transferred to 1.7 mm NMR tubes using an automated Gilson. One-dimensional (1D)-^{1}H NMR spectra and two-dimensional (2D)-^{1}H,^{13}C Heteronuclear Single Quantum Coherence Spectroscopy (HSQC).

NMR spectra were acquired using a 600 MHz Bruker Avance III spectrometer (Bruker Biospin) with an inverse cryogenic probe for 1.7 mm NMR sample tubes, fitted with a z-axis pulsed field gradient, at 300 K. Spectral widths were set to 13 and 160 ppm for the ^{1}H and ^{13}C dimensions, respectively. For the indirect (^{13}C) dimension of the 2D-^{1}H,^{13}C HSQC NMR spectra, 1228 out of 4096 (30%) data points were acquired using a non-uniform sampling scheme. ^{13}C-^{13}C splittings were enhanced 4-fold in the ^{13}C dimension. Each sample was automatically tuned, matched and then shimmed (1D-TopShim) to a DSS line width of <1 Hz before acquisition of the first spectrum. Total experiment time was ~15 min per sample for 1D-^{1}H NMR spectra and 1 h per sample for 2D-^{1}H,^{13}C HSQC NMR spectra. 1D-^{1}H NMR spectra were processed using the MATLAB-based MetaboLab software [274]. All 1D data sets were apodized using a 0.3 Hz exponential window function and zero-filled to 131,072 data points before Fourier Transformation. The chemical shift was calibrated by referencing the DSS signal to 0 ppm. 1D-^{1}H NMR spectra were manually phase corrected. Baseline correction was achieved using a spline function [274]. 1D-^{1}H-NMR spectra were exported into Bruker format for metabolite identification and concentration determination using Chenomx 7.0 (Chenomx INC). 2D-^{1}H,^{13}C HSQC NMR spectra were reconstructed using compressed sensing in the MDDNMR and NMRpipe software [275–277]. The final spectrum size was 922 real data points for the ^{1}H dimension and 16,384 real data points for the ^{13}C dimension. Analysis was performed using MetaboLab and pyGamma software was used in multiplet simulations [278]. The methyl group of lactate was used to calibrate the chemical shift based on its assignment in the human metabolome database [279].

2.5.7 Gas Chromatography-mass spectrometry (GC-MS)

These protocols were performed by Daniel A. Tennant, Paul D. Walker or Alpesh Thakker at the University of Birmingham (Birmingham, UK). Dried polar metabolites were purified as described for NMR spectroscopy. These were derivatised by incubating with 40 μL 2% methoxyamine hydrochloride (Sigma Aldrich, 226904) in pyridine (Thermo Fisher Scientific, 25104) at 60 °C for 1 h, followed by incubation with 60 μL *N*-methyl-*N*-*tert*-butyldimethylsilyltrifluoroacetamide with 1% *tert*-butyldimethylchlorosilane (MTBSTFA with 1% t-BDMCS) at 60 °C for 1 h.

GC-MS analysis was performed using an Agilent 6890GC in combination with an Agilent 5975C MS. The MS was operated under electron impact ionization at 70 eV with the source held at 230 °C and the quadrupole at 150 °C. Helium was used as the carrier gas and maintained at a flow rate of 1 mL/min. 1 μL of derivatised sample was injected (splitless) with an inlet temperature of 280 °C on to a Rxi-5MS column (Restek) The oven temperature was held at 100 °C for 1 min then increased at a rate of 5 °C/min up to a maximum temperature of 330 °C. Ions were detected using selected ion monitoring (SIM) mode as previously described [280]. MetaboliteDetector software was used to correct for the natural isotope distribution and to determine the mass isotopomer distribution (MID) [281].

2.6 Epigenome profiling

2.6.1 DNA extraction

DNA extraction was carried out using DNeasy® Blood and Tissue kit (Qiagen, Manchester, UK), according to kit protocol for use on cultured cells, using a spin-column technique. Cell culture wells had 200µL PBS added and cells were scraped and transferred to microcentrifuge tubes, followed by 20 µL PureLink proteinase K (Qiagen, Manchester, UK). Tubes were vortexed and incubated at 56 °C for 10 min. Following incubation, tubes were vortexed thoroughly for 15 s, and 20 µL RNase was added. The tubes were incubated at room temperature for 2 minutes, before 200 µL Buffer AL was added, and vortexed again. 200µL 100% ethanol (Fisher Scientific,

Loughborough, UK) was added, before mixing thoroughly again by vortexing. This mixture was transferred to a DNeasy mini spin column placed in a 2 mL collection tube. Samples were centrifuged at 6000 x g for 1 min. Flow-through was discarded. 500 µL Buffer AW1 was added, samples were centrifuged again, and flow-through discarded again. This process was repeated using Buffer AW2. Samples were then eluted in 100 µL Buffer AE. Once extracted, DNA was handled on ice and stored at -4 ^{0}C.

2.6.2 DNA quality control

DNA was quantified with the Qubit 2.0 Fluorometer (Invitrogen, Paisley UK) using DNA "broad range" standards as per manufacturer's instructions for DNA samples. DNA integrity was determined by resolving extraction products on a 1.5% agarose gel and staining with gel red (Cambridge Bioscience, Cambridge UK) to a final concentration of 0.005%.

2.6.3 Ultra-performance liquid chromatography (UPLC)

Sample preparation

2.5 µg DNA in 43 µL water, per sample, was incubated at 95 °C for 10 min. 5 µL of 10x T7 DNA polymerase reaction buffer and 10 U T7 DNA polymerase (Thermo Scientific, Renfrew, UK) was then added. The reaction was incubated overnight at 37 °C, followed by inactivation by incubation at 75 °C for 10 min. The samples were then centrifuged for 45 min at > 12,000 x g, the supernatant was aspirated and dried down using a Speed Vac and stored at -80 °C until time of analysis.

UPLC

UPLC was carried out by Dr Jimi Wills at the mass spectrometry facility, Institute for Genetic and Molecular Medicine, University of Edinburgh. After reconstitution, 10 µL polymerase-treated sample was loaded onto a HyperCarb column (Thermo Fisher Scientific, West Sussex, UK) and eluted using an Ultimate 3000 BioRS (Thermo

Fisher Scientific, West Sussex, UK) system over a 5 min gradient from 0 to 100% B, where A was 20 mM ammonium carbonate and B was acetonitrile. Data were collected using a Q Exactive (Thermo Fisher Scientific, West Sussex, UK) scanning in negative mode at 70k resolution and a scan range of 300-350 m/z.

2.6.4 Hydroxymethylcytosine DNA immunoprecipitation (hMeDIP-seq)

DNA shearing

DNA was fragmented using the Covaris® E220 Focused-ultrasonicator (Covaris, Brighton, UK). 5 µg DNA was diluted in 130 µL TE buffer and was placed into a crimp cap microTUBE (Covaris, Brighton, UK). DNA was then sonicated in the conditions shown in Table 2.11. Quality of shearing was assessed by running in a 1.5% agarose gel in 0.5x TBE buffer for 90 min at 70 volts. Acceptable shearing was deemed as fragments between 100 and 600 base pair length with a mean of 200 to 300 base pairs and a uniform appearance between sheared samples.

Peak Incident Power (W)	175
Duty Factor	10%
Cycles per burst	200
Treatment time (s)	305
Temperature (°C)	7

Table 2.8 Covaris conditions for DNA shearing

Hydroxymethylcystosine DNA immunoprecipitation (hMeDIP)

2.5µg of sheared DNA was diluted up to 450 µL in TE buffer in 1.5 mL Eppendorf® LoBind microcentrifuge tubes (Eppendorf, UK) and then denatured by placing in boiling water in a heat block for 10 min. Samples were immediately cooled on ice for 5 min. 50 µL of 10X IP buffer was then added and samples were briefly vortexed. 50 µL 'input' was then removed and placed at 40 °C until the DNA clean up step below. 1 µL of 5-Hydroxymethylcytosine (5-hmC) antibody (pAb) (Cat. No 39769; Active

Motif, CA, USA) or 1 µL Normal Rabbit IgG (Cat. No 12-370, Merck, Massachusetts, USA) was then added to the remaining sample and incubated for 2 hours at 40 °C.

40 µL of Dynabeads® Protein G for Immunoprecipitation (Invitrogen, Paisley, UK) were prewashed with 0.1% BSA in PBS for 5 min on a rotating wheel at room temperature. The beads were collected on a magnetic rack and the supernatant was removed. To ensure all supernatant was removed, the beads were briefly centrifuged at 400 x g for 5 seconds, bound to the rack and the supernatant removed again. This washing step was repeated twice. The beads were then resuspended in 40 µL of 1X IP buffer and added to the DNA/antibody mixture. This was then incubated for 1 hour at 4 °C on a rotating wheel. The beads were then collected again on the magnetic rack. The supernatant was discarded, and the tube was centrifuged again at 400 x g for 5 seconds, beads bound to the magnetic rack and residual supernatant removed. The beads were then washed three times with 1 mL of cold 1X IP buffer in the same manner as described above with BSA-PBS. The beads were then resuspended in 250 µL of digestion buffer and 20 µL of proteinase K 20 mg/ml (Roche, UK) was added. The samples were then incubated in a thermoshaker overnight at 1000 rpm and 55 °C. Following this, the tubes were spun at 400 x g for 5 seconds, the beads were then collected on a magnetic rack and the supernatant collected, containing DNA fragments enriched with 5hmC. Both the enriched fraction and input samples were then purified using the QIAquick PCR Purification Kit (Qiagen, UK) according to manufacturer's protocol, with elution into 22 µL of water.

qPCR of immunoprecipitated products

Immunoprecipitated products were analysed by qPCR to determine the efficiency of the immunoprecipitation. Regions from the genome with known high or low 5hmC levels were selected as positive and negative control regions. 10 µL of the eluted DNA from purification was diluted 1:5 in water and was used in the reaction mixture shown in Table 2.9, with primer sequences for control regions shown in Table 2.10.

A water control was used routinely. Standard curves were created using serial 1:2 dilutions of sonicated DNA. qPCR plates were centrifuged at 563 x g for 2 min prior to thermal cycling, to ensure all samples and mastermix were at the bottom of the well. The LightCycler® 480 Instrument (Roche, Sussex, UK) was used for temperature cycling (using the program shown in Table 2.11) and fluorescence quantifications. The LC480 Software (Roche) measured fluorescence throughout the experiment and data was analysed as previously described in section 2.4.4.

Reagent	**Volume (µL)**
SYBR Green Master Mix (Roche, Burgess Hill, UK	5.0
Forward Primer (100 µM)	0.02
Reverse Primer (100 µM)	0.02
Water	1.96

Table 2.9 qPCR components per well

Gene & region	**Control for 5hmC**	**Forward (5' → 3')**	**Reverse (5' → 3')**
GAPDH promoter	Negative	CGGCTACTAGCGGTTTTACG	AAGAAGATGCGGCTGACTGT
H19 genic	Positive	GATCTCGGCCCTAGTGTGAA	GTGATGTGTGAGCCTGCACT
UBIAD1 genic	Positive	CTCTTCCTCCTCCTCGTCCT	CATCCAGGAACCACAGTCCT

Table 2.10 Primer sequences used for qPCR of *hMeDIP*

Step	Temperature (°C)	Time	No. cycles
Activation	95	10 minutes	1
Denaturation	95 60	15 seconds 1 minute	70
Melting curve	95 65 75	5 seconds 1 minute Continuous (5 acquisitions/°C)	1
Cool	40	10 seconds	1

Table 2.11 LightCycler 480 thermal cycling program used with SYBR Green assays for qPCR of *hMeDIP* products

Whole genome amplification (WGA)

WGA was performed using the SeqPlex Enhanced DNA Amplification Kit (SEQXE) (Sigma, UK), according to the manufacturer's instructions. 10 µL of immunoprecipitation product and input sample were used. 15 cycles of DNA amplification were performed in a standard PCR machine (Techne Prime, Bibby Scientific, Birmingham, UK). For adaptor removal, all samples were quantified by nanodrop, the lowest DNA concentration was taken, and volume required for 2.1 µg calculated. This volume was then used for all samples in the adaptor removal step. Final elution was in 30 µL of water.

Ion Torrent Proton sequencing

Proton sequencing was outsourced to the Wellcome Trust Clinical Research Facility at the Western General Hospital, Edinburgh, UK. All kits and hardware for this procedure were from Life Technologies (Paisley, UK) unless specified.). Three biological replicates were sequenced per group. Briefly, samples were quantified using the Qubit dsDNA HS kit and approximately 100 ng of DNA was used to generate a DNA library from each sample using the Ion Xpress Plus Fragment Library Kit. During this process, DNA fragments are end repaired and then ligated to Ion specific barcode adaptors before being amplified (8 cycles) and twice purified using the

Agencourt AMPure XP PCR clean up kit (Beckman Coulter, High Wycombe, UK) which size selects fragments approximately 100–250bp in length. Libraries were then quality controlled using the Agilent Bioanalyser DNA HS kit (Agilent, Santa Clara, US) and pooled in equimolar pairs prior to template preparation using the Ion PI™ Hi-Q™ OT2 200 Kit and sequencing on the Ion Torrent semiconductor sequencer using the Ion PI™ Hi-Q™ Sequencing kit and an Ion PI™ Chip Kit v3. For consistency, each sample was sequenced on a PI chip with its own input.

2.7 Bioinformatic and statistical analyses

2.7.1 Bioinformatic analysis of RNA-seq data

RNA-seq analysis

The generated FASTQ files were trimmed to remove adapters, using Trimmomatic (version 0.36) [282], before performing quality control with FastQC (version 0.11.4) [283]. Alignment was performed against the *Homo sapiens* GRCh19 assembly. The assembly was first indexed using STAR (version 2.5.1b) before mapping trimmed reads, using STAR (version 2.5.1b) in paired-end mode with default behaviour [284]. Duplicate reads were removed using Picard (version 2.7.11) [285], before using featureCounts to generate raw read counts for each gene. Differential gene expression (DEG) analysis was performed using DESeq2 [286]. Heatmaps were generated with Heatmapper [287]. Pathway enrichment analysis was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) function [288–290] of the Database for Annotation, Visualization, and Integrated Discovery (DAVID) [291,292]. Transcriptomic sequencing data have been deposited on the GEO repository (accession number GSE138052).

3' RNA-seq analysis

Data from 3' RNA-sequencing were analysed using the BlueBee platform (Lexogen, Netherlands), which generated differential expression and principal component analyses (PCA). Pathway enrichment analysis was performed using the Gene Ontology (GO) Biological Processes function [288–290] of DAVID [291,292].

2.7.2 hMeDIP-seq analysis

Sequencing generated reads with a mean read length of 137-147 base pairs and 21,130,039 - 31,693,844 reads per sample. Reads were aligned to the hg19 genome using Torrent Suite v5.2.0. Aligned reads were sorted using SAMtools, before calling peaks using MACS2 (v. 2.1.1) -f BAM --broad --broad-cutoff 0.05 -B -g hs, over corresponding inputs [293]. To detect differentially hydroxymethylated regions (DHRs), I used Diffbind with DESeq2 [294]. For Diffbind analysis, data were normalised to a pooled input for each group and an IgG control. Differentially hydroxymethylated regions were assigned to genes and other genomic features using the HOMER (v. 4.8; hg19) annotatePeaks tools [295]. Regions of interest were identified from the hmeDIP-sequencing dataset. Data are available through the Gene Expression Omnibus (GSE144955). Sliding window profiles and heatmaps were generated using deepTools [296], using the plotProfile and plotHeatmap functions, respectively, with blacklisted regions subtracted.

2.7.3 Statistical testing of null hypotheses

Statistical analysis was performed using GraphPad Prism 8 software. Datasets were routinely tested for normality of distribution using the Shapiro-Wilks test. When comparing two groups, where data were normally distributed, an unpaired Student's t-test was used for analysis. Where data were not normally distributed, a non-parametric Mann Whitney U test was used. Comparisons between two groups with repeated measures (e.g. different isotopomers of the same metabolite) were made using a two-way analysis of variance (ANOVA), with Sidak post-hoc testing to account for the false discovery rate. All values are represented as mean ± SD with $p < 0.05$ classed as significant.

Chapter 3: Characterising the transcriptome of a human pluripotent stem cell model of NAFLD

3.1 Introduction

In the earliest stage of NAFLD, TG accumulation in hepatocytes leads to the development of hepatic steatosis [17], which is characterised by the development of macrovesicular steatosis, whereby TGs are stored in large lipid droplets [297]. Understanding the changes in mRNA expression associated with development of these lipid droplets may be informative for the development of therapeutics in the future.

Accurate *in* vitro modelling of human NAFLD is challenging, and to date many groups have focused on the use of whole tissues from human or mouse, HCC cell lines, or primary hepatocytes [42,129,298,299]. However, the use of whole tissue for transcriptome analysis is confounded by the presence of heterogeneous cell populations, each of which may have a different transcriptional profile. A limitation of using primary cells is that they rapidly lose their phenotype and display altered transcriptional signatures on purification [300,301]. I therefore set out to further characterise a human-relevant in vitro model of NAFLD [128], that overcomes the limitations of other models and determine the impact of hepatic steatosis on the transcriptome. To address this, I performed high content analysis microscopy to analyse patterns of lipid accumulation in LPO-treated HLCs, before performing an in-depth bioinformatic analysis of the transcriptome in these cells and comparing this with publicly available data from NAFLD patients.

3.1.1 Hypothesis

Treatment of HLCs with LPO induces similar lipid accumulation and transcriptional patterns to those observed in human disease making this a suitable model for analysis of NAFLD-associated metabolic and epigenetic dysfunction.

3.1.2 Aims

1. To determine whether HLCs differentiated on a Laminin-521 extracellular matrix develop a NAFLD-like phenotype upon challenge with high energy substrates.
2. To assess patterns of lipid accumulation in LPO-treated HLCs, and whether they mirror human pathology.
3. To characterise the transcriptomic changes that occur in LPO-treated HLCs and comparing these with published human data.

3.2 Materials and methods

3.2.1 Differentiation and maturation of hPSC-derived hepatocyte-like cells

To differentiate human female H9 cells to HLCs, I used a previously defined protocol, as outlined in **Figure 3.1** and fully described in Chapter 2, section 2.2. H9s were primed to a definitive endoderm using a combination of WNT3A and Activin A, with daily media changes for 3 days. Next, I induced hepatoblast specification by culturing endodermal cells in media supplemented with Knock-Out Serum Replacement and DMSO. Hepatoblast specification continued for 5 days, with media changes every other day. To mature hepatoblasts to a hepatocyte-like phenotype, cells were then cultured in media containing oncostatin M, hepatocyte growth factor, and hydrocortisone. Maturation continued for 9 days, with media changes every other day.

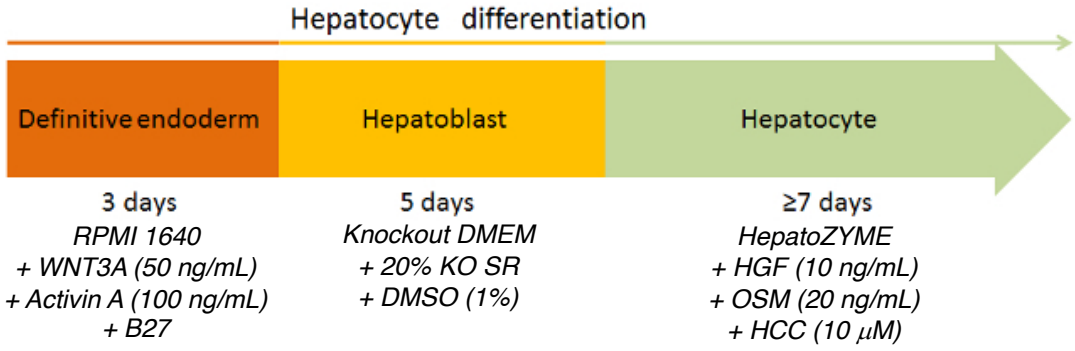


Figure 3.1. Schematic of the hepatocyte-like cell differentiation protocol. This 3-stage differentiation protocol differentiates hPSCs from seeded single cells (day 0) to a definitive endoderm (day 3), hepatoblast specification (day 8), forming mature HLCs (day 17). RPMI = Roswell Park Memorial Institute (RPMI) 1640 medium; B27 = B27 medium supplement; Knockout DMEM = Knockout Dulbecco's Modified Eagle Medium; KO SR = knockout serum replacement; HGF = hepatocyte growth factor; OSM = oncostatin M; HCC = hydrocortisone.

HLCs were cultured in a 96-well format for measurements of lipid accumulation and in a 6-well format for all other analyses. The luciferase-based assay to measure CYP3A4 activity is described full in Chapter 2, section 2.3.2. Lipid accumulation was

induced as previously described [128]. Briefly, at day 17, HLCs were exposed to a cocktail of sodium l-lactate (L; 10mM), sodium pyruvate (P; 1 mM) and octanoic acid (O; 2 mM) (Sigma, Gillingham, UK) for a period of 48 h.

3.2.2 High content analysis microscopy

Cells were stained with a cell painter assay, adapted from Lyall *et al* and Bray *et al* [128,302]. Cells were fixed with 50 µL/well 4% (wt/vol) paraformaldehyde (Electron Microscopy Sciences, 15710-S) for 15 minutes at room temperature. For permeabilisation, cells were incubated in 0.1% Triton X-100 (Sigma-Aldrich, T8787) in PBS for 15 minutes at room temperature. For lipid droplet analysis, cells were then stained with a combination of NucBlue Live ReadyProbes® Reagent (2 drops/mL) (Molecular Probes, R37605), HCS CellMask™ Red (2 µL/10 mL) (Invitrogen, H32712), and BODIPY™ 493/503 (1:1000) (Life Sciences, D3922), as per the manufacturer's instructions. Following staining, images were acquired using an Operetta High Content Analysis microscope (Perkin Elmer, Buckinghamshire, UK). Lipid droplet morphology was analysed as described in Chapter 2, section 2.3.4.

3.2.3 Gene expression

Total RNA was extracted from HLCs using the Monarch® Total RNA Miniprep Kit (New England BioLabs, T2010). cDNA was generated using the High Capacity cDNA Reverse Transcriptase Kit (Applied Biosystems, 4368814). A master mix was prepared using PerfeCTa FastMix II (Quanta Biosciences, Inc., 95118-250). cDNA was amplified and quantified using the Universal Probe Library (Roche, Burgess Hill, UK) system on a Roche LightCycler 480 (Roche Diagnostics Ltd, Switzerland). Gene expression was measured for analysis of lipid droplet-associated markers and validation of transcriptomic analysis. Details of primers and Universal Probe Library probes (Roche) can be found in Chapter 2, Table 2.6.

3.2.4 Transcriptomic and pathway analysis

RNA library preparation and sequencing were performed by BGI (Beijing, China). The sequencing and bioinformatic pipelines are fully described in Chapter 2, section 2.4.7.

3.2.5 Statistics

All statistical analyses were performed using Graph Prism Version 8.0 for Windows or macOS, GraphPad Software, La Jolla California USA, www.graphpad.com. Normality of data distribution was measured using the Shapiro-Wilks test. Where indicated, data were analysed by unpaired Student's t-test, Mann-Whitney test or two-way analysis of variance (ANOVA). Data were considered to be significant where $p < 0.05$. For transcriptomic analyses, changes in mRNA expression were considered to be significant where the adjusted p value <0.05.

3.3 Results

3.3.1 Characterisation of human pluripotent stem cells (hPSCs)

Standardised and reproducible hPSC culture requires a number of quality control assessments to ensure maintained pluripotency of the cells, including morphological validation and examination of pluripotency and cell surface markers. For the studies described throughout this thesis, I used a female H9 hPSC line, which were cultured on Laminin-521-coated plates. To minimise spontaneous differentiation, cells were passaged regularly. Cell colonies were tightly packed and had a high nucleus to cytoplasm ratio (**Figure 3.2**).

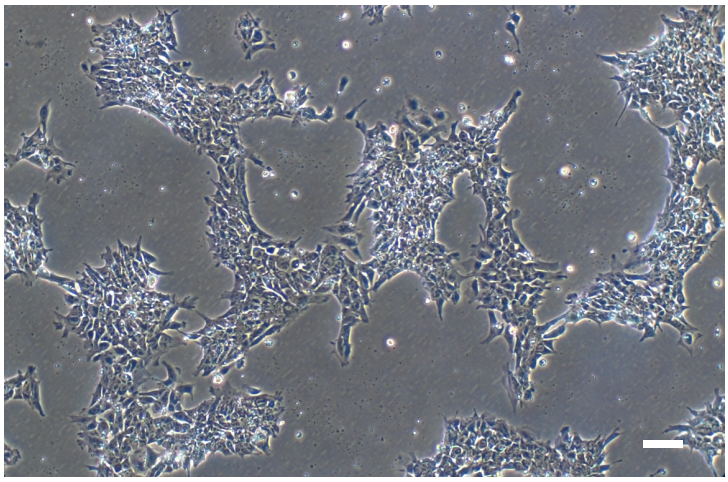


Figure 3.2. hPSC morphology in culture. Representative image of hPSCs, cultured in mTeSR medium, on Laminin-521-coated 6-well plates. Imaged H9 cells are at passage 40. Scale bar = 100 μm.

Using immunohistochemistry, I showed that H9 cells, in the undifferentiated state, express NANOG, marking their pluripotent capacity (**Figure 3.3**). To assess the presence of pluripotent cell surface markers, I performed flow cytometry and measured SSEA-1, SSEA-4, TRA-1-60 and TRA-1-81. While H9s were negative for SSEA-1, they were positive for SSEA-4, TRA-1-60 and TRA-1-81, further supporting that these cells have pluripotent capacity (**Figure 3.4**). The capacity of these cells to

differentiate to hepatocyte-like cells demonstrates that they are also capable of forming an endodermal layer.

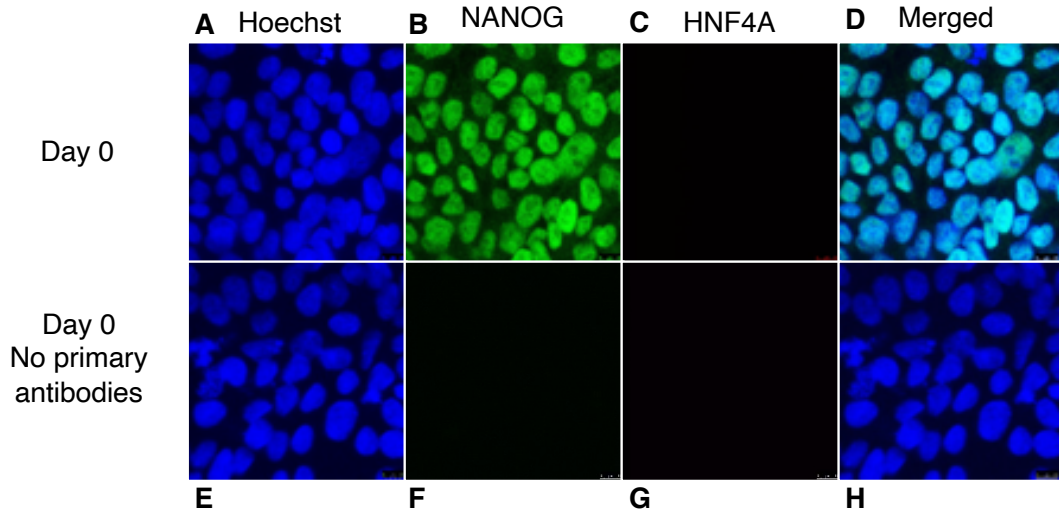


Figure 3.3. Pluripotent marker expression in hPSCs. Pluripotent markers were detected by immunofluorescence in H9 hPSCs. (**A**) Hoechst dye was used to stain nuclei. (**B-C**) Staining was performed against a marker of pluripotency (NANOG) and a hepatocyte marker (HNF4A). (**D**) Images were merged to demonstrate localisation of NANOG to the cell nucleus in H9 cells. (**E-H**) Staining protocol was repeated in the absence of primary antibodies to ensure lack of non-specific binding by secondary antibodies. Images were taken at 40X magnification.

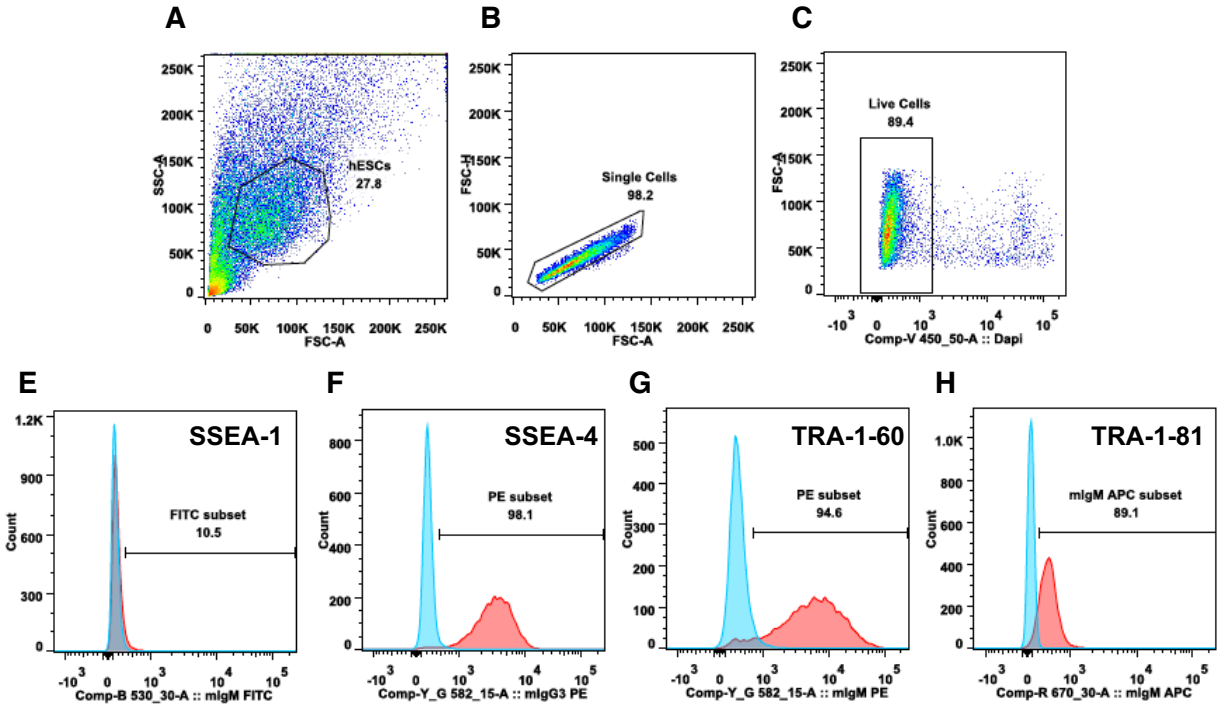


Figure 3.4. H9 cells express cell surface markers associated with pluripotency, which were detected by flow cytometry. Flow cytometry was used to determine the presence of pluripotency-associated cell surface markers in H9 cells. (**A-C**) Schematic of flow cytometry analysis methodology. H9s were negative for the expression of a marker of differentiation (**E**) SSEA-1. In contrast, H9s expressed the pluripotency markers (**F**) SSEA-4, (**G**) TRA-1-60, and (**H**) TRA-1-81. Figure panels show a representative distribution. *n* = 4 biological replicates.

To fully assess differentiation of hPSCs to HLCs, I performed RT-qPCR for pluripotency and hepatic markers, as well as measuring hepatocyte activity. NANOG mRNA expression was measured at each stage of differentiation, to determine whether cells were losing their pluripotent capacity (**Figure 3.5**). There was a decrease in NANOG expression throughout the differentiation process, which became undetectable once HLCs reached the end of the maturation period. In contrast, the hepatic markers ALB and HNF4A were undetectable in the undifferentiated H9s but were robustly expressed once cells reached the end of the maturation period. At this stage, CYP3A4 activity was analysed at the end of the differentiation process, to determine whether HLCs were displaying functional hepatocyte activity (**Figure 3.6**). Once fully differentiated, HLCs displayed robust CYP3A4 activity, in line with previous studies [128], and this was not diminished by induction of intracellular lipid accumulation, indicating that cells retain their hepatocyte-like phenotype upon challenge.

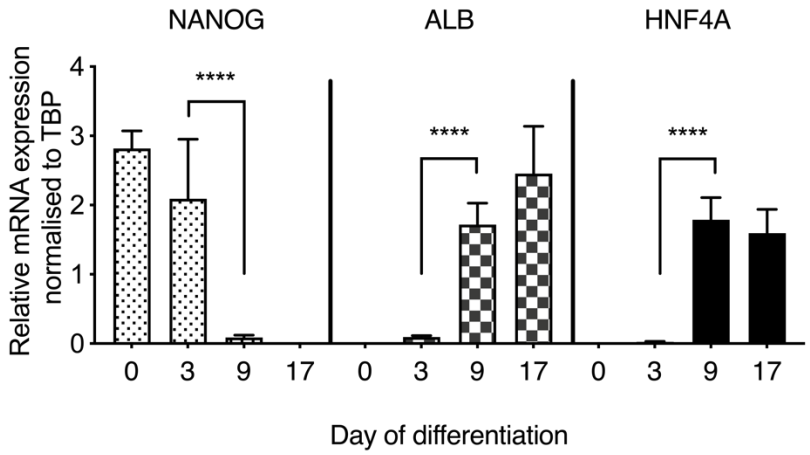


Figure 3.5. Analysis of pluripotency and hepatic markers in HLCs during the differentiation process. RT-qPCR was used to measure mRNA expression of pluripotency (NANOG) and hepatic (ALB and HNF4A) markers. For each time point *n* = 3 biological replicates per gene. Data were analysed by two-way ANOVA with Sidak post-hoc testing and expressed as mean ± SD. ****$p<0.0001$. Figure taken from Sinton *et al* (2020) [303].

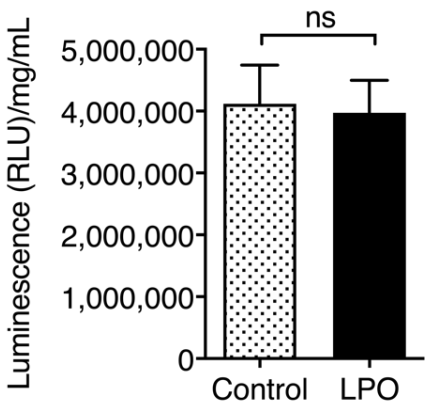


Figure 3.6. Functional analysis of hepatocyte activity in fully differentiated HLCs. CYP3A4 activity was measured to determine levels of functional hepatocyte activity. For each group *n* = 8 biological replicates. Data were analysed using a two-tailed Student t-test and expressed as mean ± SD. Figure taken from Sinton *et al* (2020) [303].

3.3.2 LPO treatment induces macrovesicular steatosis in HLCs

It was previously reported that the high energy substrate cocktail LPO promotes intracellular lipid accumulation in HLCs [128] but it was unknown whether it promotes lipid droplet biogenesis or increased accumulation of lipids within existing droplets. To further assess this, mRNA expression of PLIN2, PLIN4 and PLIN5 was measured (**Figure 3.7A-C**), which showed significant increases in association with LPO treatment. I then performed lipid staining with BODIPY dye and assessed lipid droplet hyperplasia and hypertrophy. Representative images of control and LPO-treated cells indicated a perturbation of lipid accumulation (**Figure 3.8A-B**). Further analysis of lipid droplets using a high content microscopy system showed that LPO exposure was not associated with a change in the number of lipid droplets (**Figure 3.8C**) but did lead to an increase in the size of intracellular lipid droplets (**Figure 3.8D**), consistent with the development of macrovesicular steatosis.

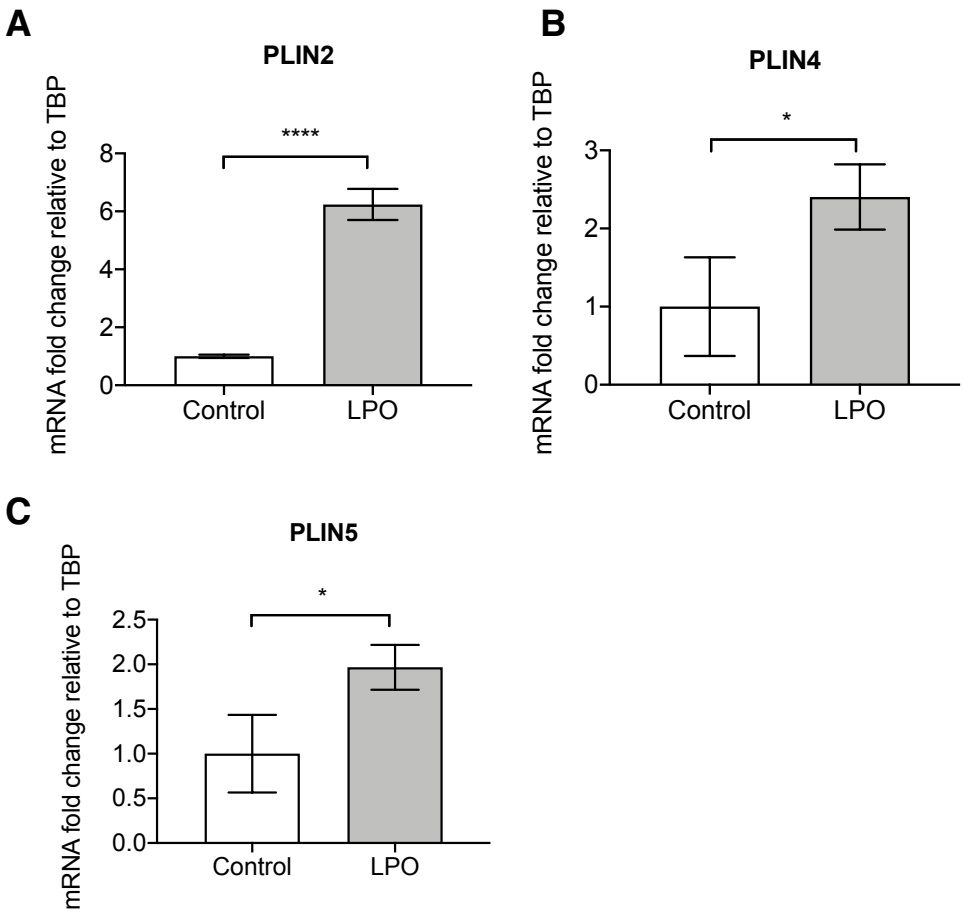


Figure 3.7. RT-qPCR analysis of genes associated with lipid droplet membranes. RT-qPCR was used to measure mRNA expression of genes typically associated with formation of lipid droplet membranes. Expression of PLIN2 (**A**), PLIN4 (**B**) and PLIN5 (**C**) was measured. For each group $n = 3$ biological replicates. Data were analysed using a two-tailed Student t-test and expressed as mean ± SD. *$p<0.05$, ****$p<0.0001$. PLIN = Perilipin. Data adapted from Sinton *et al* (2020) [303].

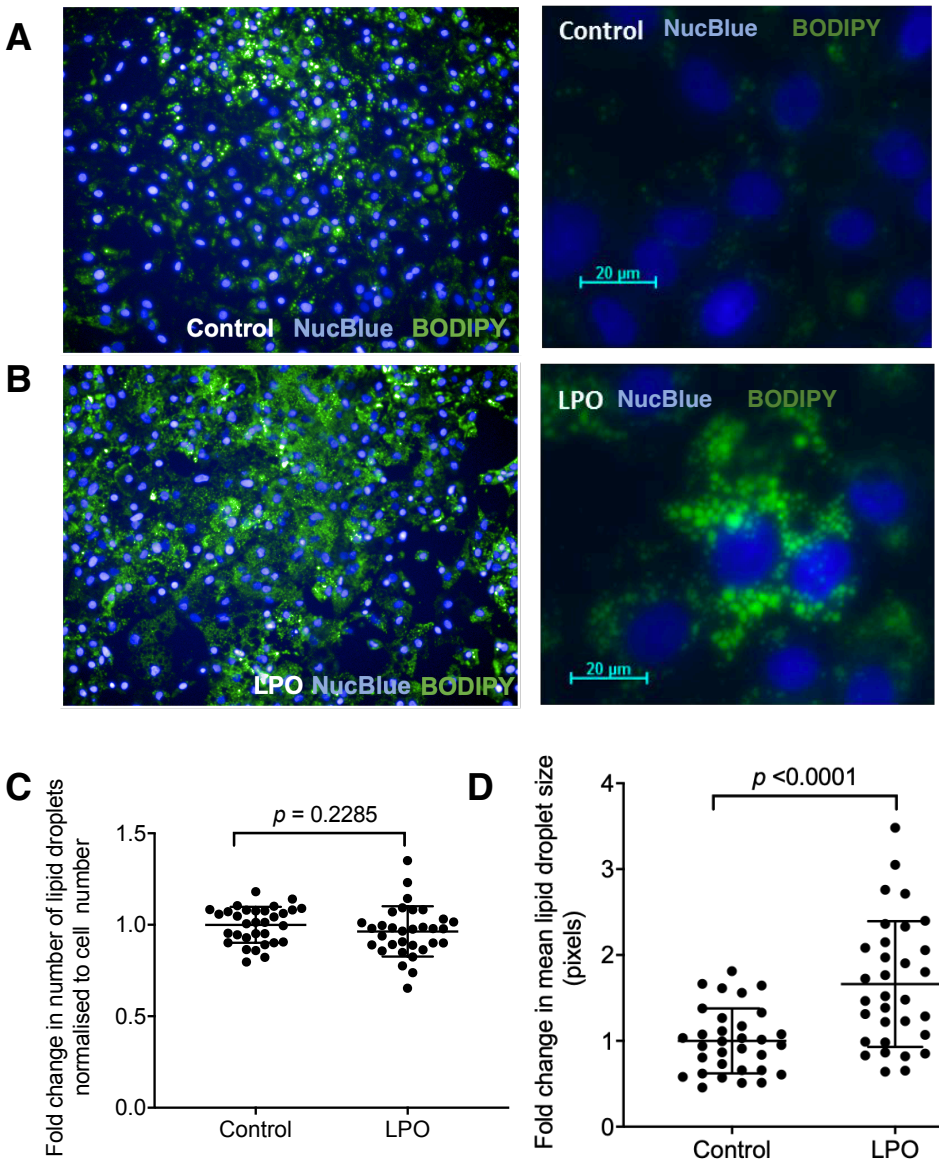


Figure 3.8. High content analysis of lipid droplets in LPO-treated HLCs. Intracellular lipid droplet number and size were measured by high content analysis microscopy. Hoechst dye (NucBlue) and BODIPY were used to stain nuclei or neutral lipids, respectively. Representative images of lipid accumulation in control (**A**) and LPO-treated (**B**) groups; 10x or 40x magnification. (**C**) Number of lipid droplets do not increase, (**D**) but intracellular lipid droplets increase in size (n = 32 biological replicates/group). Data were analysed using two-tailed Student t-test and expressed as mean ± SD. LPO = Lactate, Pyruvate, Octanoate. Figure adapted from Sinton *et al* (2020) [303].

3.3.3 Macrovesicular steatosis in HLCs is associated with wide-scale alterations of the transcriptome

To examine the impact of LPO treatment on the transcriptome, mRNA sequencing was performed to a depth >60 million reads. Differential expression analysis identified 853 downregulated and 826 upregulated genes, with a Log2 fold change cut-off >1.5 (**Supplementary file MCS1 – DEGs in LPO-treated HLCs**). The results of the RNA-seq pipeline were validated by RT-qPCR analysis of key genes (**Figure 3.9**).

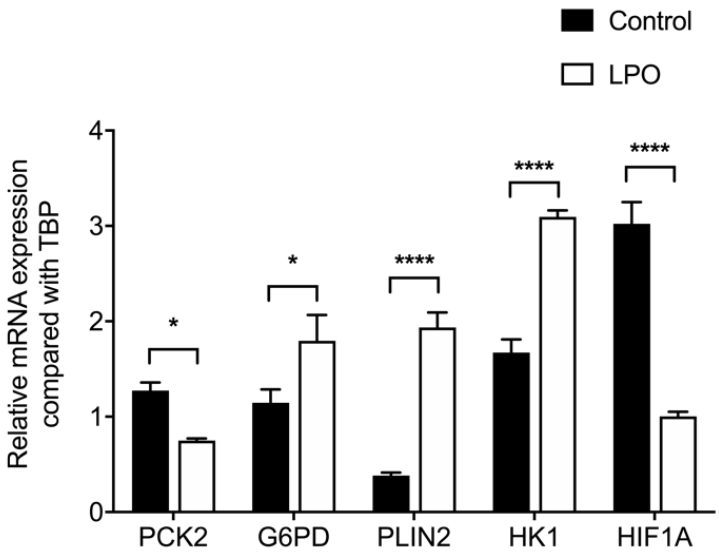


Figure 3.9. Validation of RNA-seq analysis by RT-qPCR. Key genes were selected from the generated RNA-seq analysis and measured by RT-qPCR to confirm the analysis output. $N = 4$ biological replicates/group. Data were analysed using two-tailed Student t-test and expressed as mean ± SD. *p<0.05, ****p<0.0001. Figure adapted from Sinton *et al* (2020) [304].

Principal component analysis (PCA) demonstrated that replicates within each of the control and LPO group clustered, showing minimal variation between replicates (**Figure 3.10**). Heatmap analysis of the top 1,000 dysregulated genes highlights the widescale transcriptomic changes that occur following LPO treatment (**Figure 3.11**). Furthermore, when applying statistical cut-offs of a Log_2 fold change >1.5 and adjusted p value <0.05, wide-scale transcriptional changes are still present in LPO-treated HLCs (**Figure 3.12**). Taken together, these findings demonstrate that LPO

treatment induces a transcriptional pattern that is distinct from control HLCs, and that this is consistent between biological replicates.

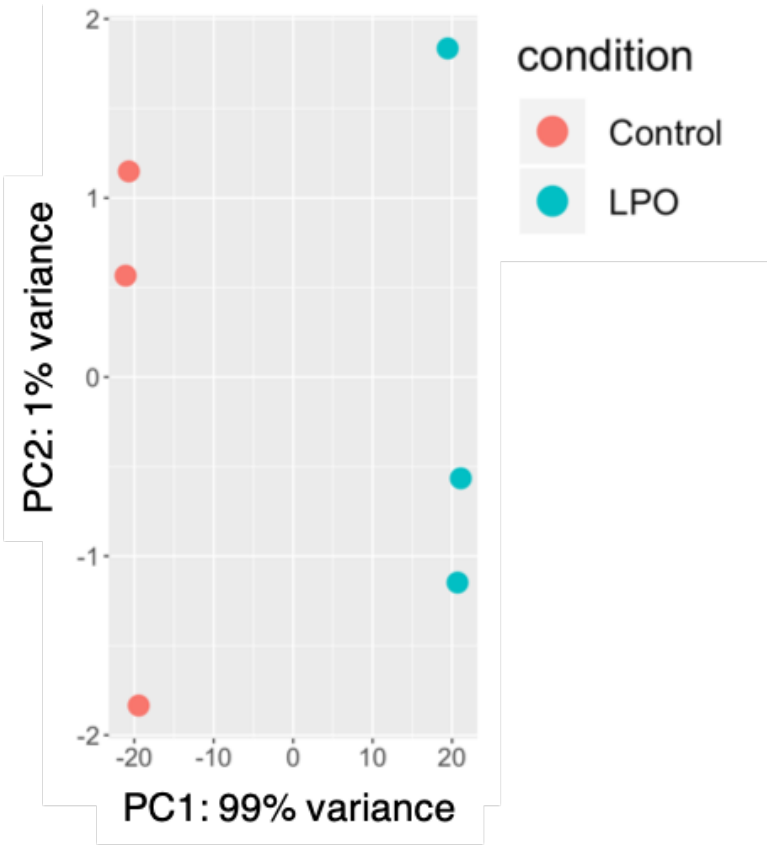


Figure 3.10. Principal component analysis of transcriptomic data reveals clustering of replicates within groups following treatment of HLCs with LPO. PCA plot demonstrating differential clustering of each biological replicate from both control and LPO-treated HLCs. For each group *n* = 3 biological replicates. Figure adapted from Sinton *et al* (2020) [304].

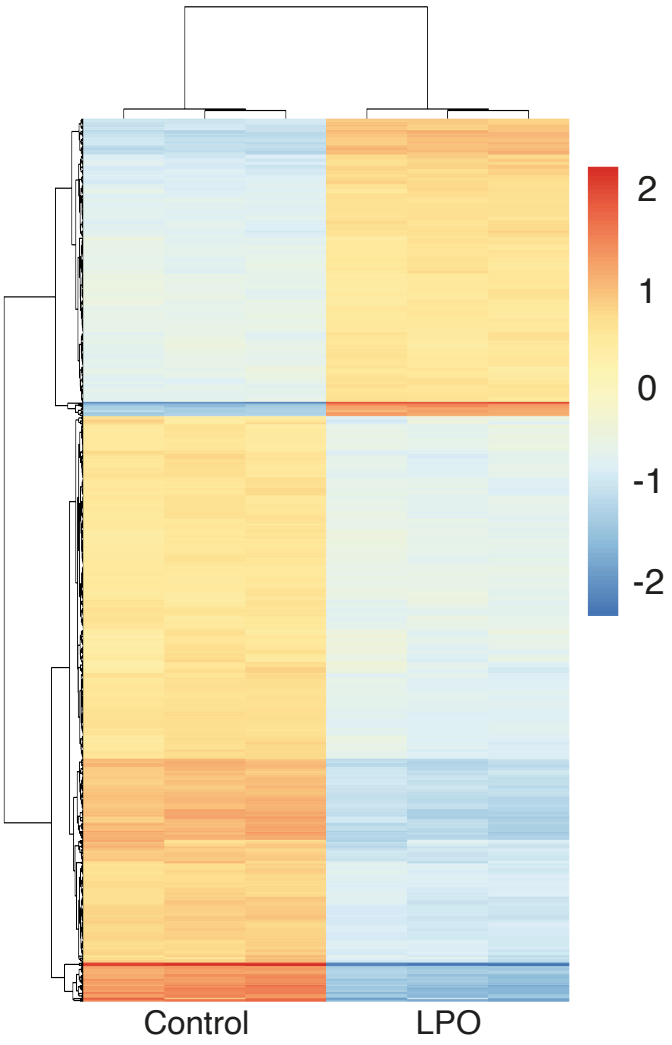


Figure 3.11. Heat map analysis of transcriptomic data reveals wide-spread transcriptomic changes following treatment of HLCs with LPO. Heap map representing the top 1,000 differentially expressed genes when comparing control and LPO-treated HLCs. Genes are clustered by Euclidean distance. For each group *n* = 3 biological replicates. Figure adapted from Sinton *et al* (2020) [304].

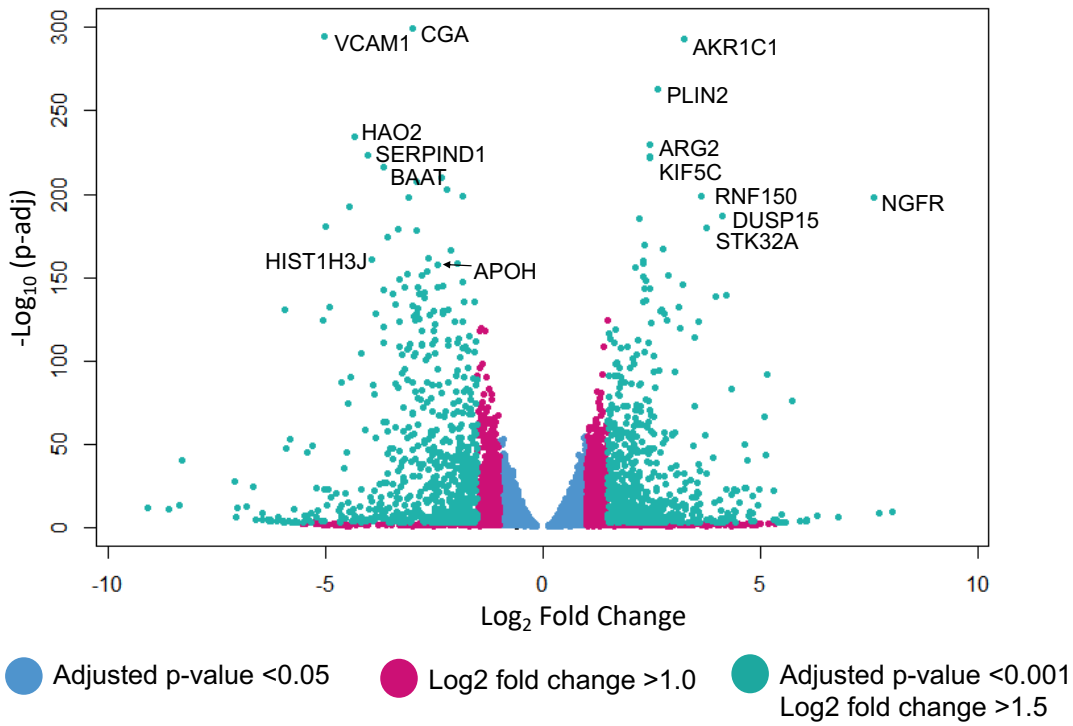


Figure 3.12. Volcano plot of transcriptomic data reveals wide-spread transcriptomic changes following treatment of HLCs with LPO. Representation of differential mRNA expression in control and treatment groups. Non-significant expression is represented in either blue (adjusted *p* value >0.05) or purple (Log_2 fold change <1.0). For each group *n* = 3 biological replicates. Figure adapted from Sinton *et al* (2020) [304].

Within the differentially expressed genes identified in this analysis, a number of genes that are typically associated with NAFLD were dysregulated upon induction of macrovesicular steatosis (**Figure 3.13**). Upregulated genes included those involved in lipid metabolism (PPARGC1A, PPARG, PPARA), lipid storage (PLIN2), glucose transport (SLC2A4, which encodes the glucose transporter type 4 GLUT4), and the pentose phosphate pathway (G6PD). In contrast, downregulated NAFLD-associated genes included those involved in cholesterol and bile acid regulation (CYP7A1), insulin resistance (NR0B2), ketogenesis (HMGCS1 and HMGCS2) and fatty acid synthesis (FASN).

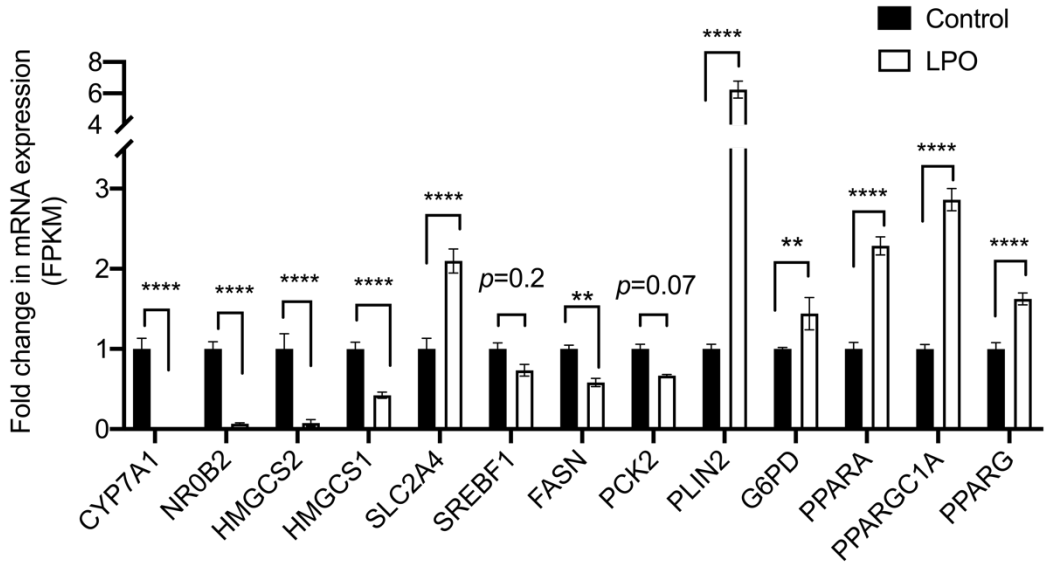


Figure 3.13. Dysregulation of NAFLD associated genes. Differential expression analysis revealed dysregulation of a number of genes that are typically associated with NAFLD. For each group n = 3 biological replicates. Data presented here represent a subset of genes from the RNA-seq dataset. Data were analysed by two-way ANOVA with Sidak post-hoc testing and expressed as mean ± SD. $^{**}p<0.01$, $^{****}p<0.0001$. Figure adapted from Sinton *et al* (2020) [304].

3.3.4 LPO treatment in HLCs disrupts pathways related to gene regulation, lipid metabolism and mitochondrial respiration

I mapped genes with a log2 fold change cut-off >1.5 to the KEGG pathway database, identifying a number of enriched pathways. Upregulated pathways predominantly comprised signalling pathways, including hsa04725: cholinergic synapse, hsa04726: serotonergic synapse, and hsa04020: calcium signalling, and has04014: Ras signalling (**Figure 3.14A**). These pathways contained genes encoding adenylate cyclases, potassium inwardly rectifying channels, G proteins, and fibroblast growth factors.

Downregulated KEGG pathways included several pathways (hsa05322: systemic lupus erythematosus, hsa05034: alcoholism, and hsa05202: transcriptional misregulation in cancer) containing a substantial number of genes encoding the histone structural units H1, H2A/B, H3 and H4 (**Figure 3.14B**; Chapter 8, Appendix 1). In addition, there was downregulation of the pathways related to cholesterol (hsa00140: steroid hormone biosynthesis) and drug metabolism (hsa00982: drug metabolism – cytochrome P450, hsa00983: drug metabolism – other enzymes, and hsa00980: metabolism of xenobiotics by cytochrome P450). The cholesterol metabolism pathway showed downregulation of genes encoding hydroxysteroid dehydrogenases, cytochrome P450, and UDP-glucuronosyltransferases. Genes encoding cytochrome P450, and UDP-glucuronosyltransferases were also identified in drug metabolism pathways.

When specifically comparing these transcriptomic data with the KEGG terms TCA Cycle and Oxidative Phosphorylation, there was extensive disruption within these pathways (**Figure 3.15**). Within the TCA Cycle pathway there were low, but significant, levels of transcriptomic disruption in the majority of genes encoding enzymes that catalyse metabolite interconversion (Chapter 8, Appendix 2). Analysis of the Oxidative Phosphorylation pathway revealed an overall downregulation of transcription of genes encoding components of respiratory complexes I (ND1, ND2, ND4L, NDUFS2, NDUFV1, NDUFA10, NDUFB2 NDUFA2, NDUFB10) and IV (COX1, COX4I1, COX8A, COX6B1), as well as ATP synthase (ATP6V1A, ATP5MF, ATP6V0D1, ATP6V1E1, ATP6, APT5MC2, ATP5MC1) (Chapter 8, Appendix 3). Previously, I observed decreased mitochondrial maximal respiration and spare reserve capacity in HLCs with increased intracellular lipid accumulation [303]. Taken together, these data suggest that NAFLD is associated with disruption of both the TCA cycle and ETC.

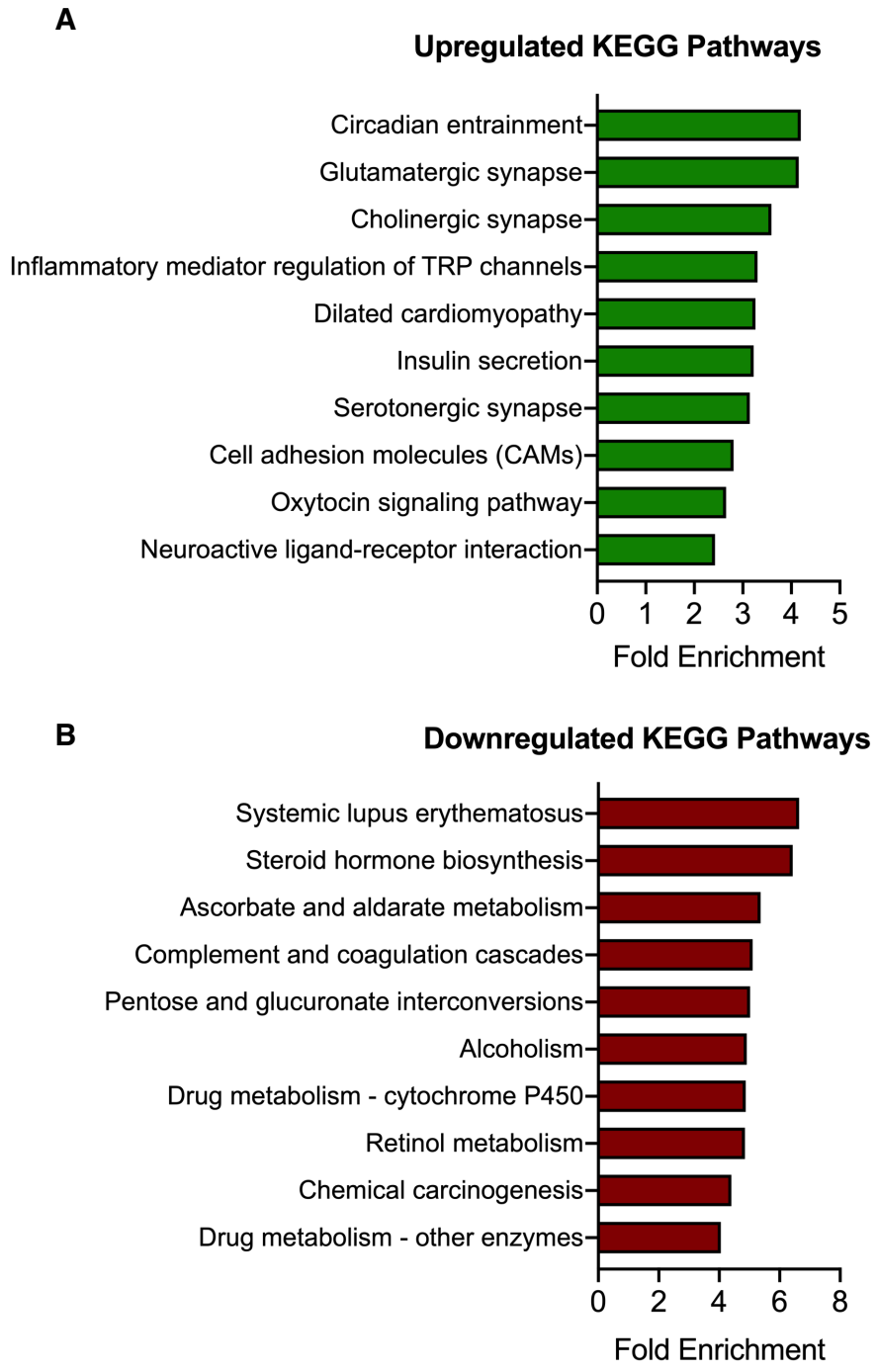


Figure 3.14. KEGG Pathway analysis. All genes with a Log_2 fold change > 1.5 were analysed with DAVID. The top 10 pathways for up- and downregulated genes are represented here. Details are located in Chapter 8, Appendix 1. Figure adapted from Sinton *et al* (2020) [304].

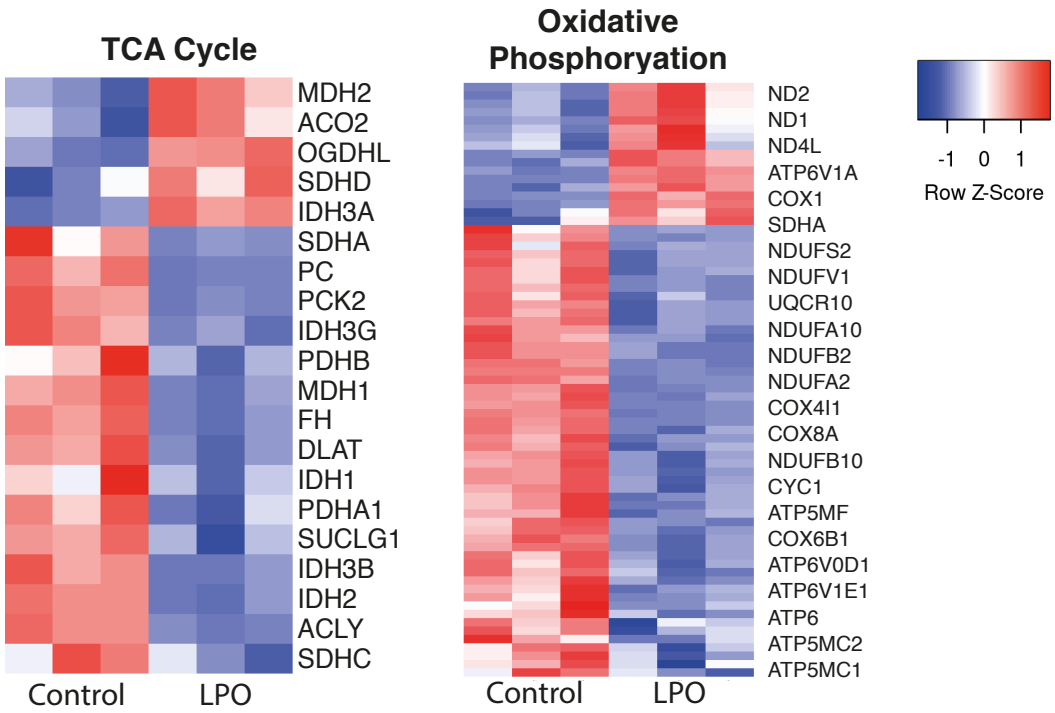


Figure 3.15. Examination of the TCA Cycle and Oxidative Phosphorylation KEGG pathways. Heatmap representing genes identified in the TCA Cycle and Oxidative Phosphorylation KEGG pathways. Gene symbols are representative, and a full list is located in Chapter 8, appendices 2-3. Figure adapted from Sinton *et al* (2020) [304].

3.3.5 Analysis of transcriptomic patterns in human NAFLD confirms disease heterogeneity

As previously stated, NAFLD is a challenging disease to study in humans, due to the spectrum of pathologies it encompasses and the difficulty in obtaining suitable control samples. A recent study from Suppli *et al* (2019) obtained hepatic tissue samples from healthy adult male and female patients, alongside samples from patients with NAFLD [298]. To specifically compare steatotic HLCs with the transcriptome of NAFLD patients, I downloaded the publicly available RNA-seq datasets from GEO (accession number GSM3615293) and ran the RNA-seq pipeline again, this time comparing human control with disease samples in both males and females. The PCA revealed that there was no clustering within groups, reflecting the heterogeneity of NAFLD

(**Figure 3.16**). Further to this, pathway enrichment analysis of these data, using differential expression cut-offs of Log_2 fold change >1.5 and adjusted *p* value <0.05 did not reveal any significantly enriched pathways. These findings precluded any meaningful comparison between the HLC- and human-based expression datasets.

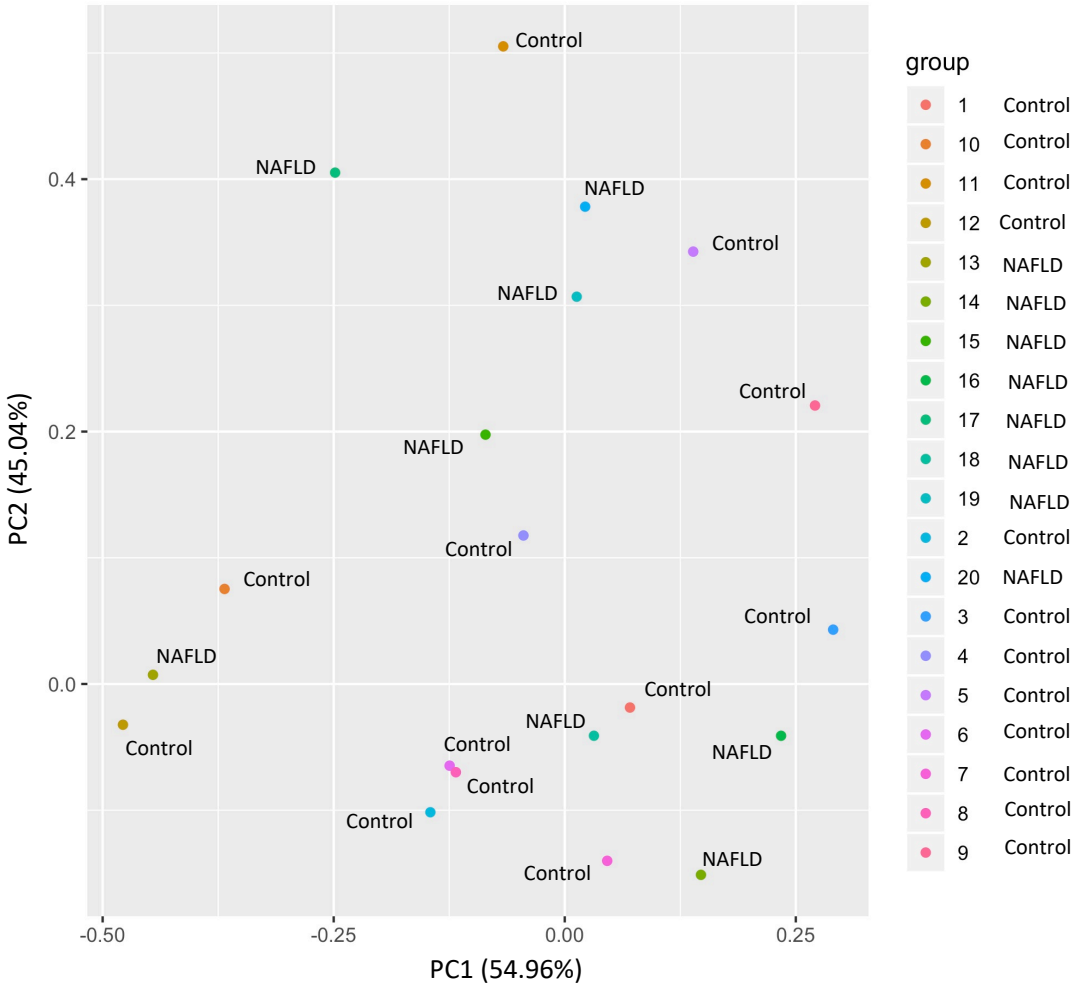


Figure 3.16. PCA plot comparing human healthy control and NAFLD transcriptomic signatures. The analysis indicates no clustering within groups, reflecting heterogeneity within the samples analysed.

3.4 Discussion

NAFLD is a challenging disease to study in humans, due to the difficulty of obtaining tissues that are not compromised by confounding conditions. Primary cells for studies are generally obtained from organs unsuitable for transplant [305] or non-affected tissue adjacent to tumours [306], potentially confounding analyses. These cells are technically challenging to purify and rapidly lose their phenotype *in vitro* [300], and adaptation to cell culture is demonstrated to alter their transcriptional profiles, which begin to resemble diseased liver tissue [301]. The Drake lab previously described a pluripotent stem cell-based model of NAFLD, where array-based sequencing revealed many transcriptomic features in common with human disease [128]. Here, I used a model that was further refined, whereby cells were cultured and differentiated on a Laminin-521 matrix instead of Matrigel. This change to the differentiation process was previously found to induce a more robust hepatocyte phenotype, with improved cell polarity and CYP4A4 activity, as well as reducing batch variation between differentiations [267].

3.4.1 LPO treatment induces macrovesicular steatosis in HLCs, mirroring patterns of lipid accumulation in human NAFLD

Following refinement of this model, genes related to lipid droplet formation were measured, and patterns of lipid accumulation were assessed. The observed changes in expression of genes associated with lipid droplet membrane formation (PLIN2 and 5) are in agreement with findings from human studies: intracellular lipid accumulation is associated with upregulation of PLIN2 in human hepatocytes [307] and immunohistochemistry staining of biopsies from NAFLD patients shows increased accumulation of PLIN2 at the surface of macrovesicular lipid droplets [308]. In adipocytes, PLIN4 localises to the surface of unilocular lipid droplets, and may participate in their early development. Although there are no human studies describing perturbed expression of PLIN4 in association with NAFLD, knockout of this gene protects against hepatic steatosis in mice [309]. Furthermore, PLIN4 knockout diminishes the accumulation of triacylglycerol in lipid droplets and reduces the expression of PLIN5 in mouse cardiac tissue. This interplay may be important for the

pathology of NAFLD as *in vitro* studies demonstrate that PLIN5 blocks lipid droplet lipolysis in hepatocytes. Therefore, increases in PLIN5 expression may also contribute to the development of macrovesicular steatosis [310].

To further understand patterns of intracellular lipid accumulations in HLCs, cells were stained with BODIPY 493/503 and examined using HCA microscopy. Whilst the number of intracellular lipid droplets did not increase following LPO treatment, their size increased, indicating the development of macrovesicular steatosis. NAFLD is typically associated with the development of macrovesicular steatosis, whereby TGs accumulate in single large or multiple medium-sized cytoplasmic lipid droplets [297,311]. Alongside analyses of PLIN2/4/5 expression patterns, these HCA data strongly indicate that this model recapitulates the macrovesicular steatosis observed in humans with NAFLD.

3.4.2 Macrovesicular steatosis is associated with widescale transcriptomic changes

Following refinement of the model, deep sequencing was performed to further understand how steatotic hepatocytes rewire their transcriptional profiles in response to intracellular lipid accumulation. Analysis of publicly available human NAFLD data [298] revealed extensive transcriptomic heterogeneity, precluding comparison with the transcriptomic dataset generated from steatotic HLCs. As NAFLD comprises a spectrum of pathologies, it is possible that samples were taken from patients at different disease stages. It is also possible that samples were taken from different locations in the liver, which may also impact analysis. Furthermore, it is possible that analyses were confounded by the use of bulk tissue analysis. These factors demonstrate the necessity of laboratory models, such as the one described here, which are reproducible and not excessively prone to batch variation

Within the steatotic HLC transcriptomic dataset, I observed changes in the expression of a number of genes that were consistent with previous observations in humans

[298]. Pathway enrichment analysis allowed me to further interrogate transcriptional rewiring in steatotic HLCs and revealed disruption related to multiple metabolic pathways. This included pathways related to steroid hormone biosynthesis and gene regulation, each of which was downregulated. This is in line with other studies of NAFLD showing downregulation of many of the genes included in these pathways [298,312–318]. Whilst it has been documented that 17β-hydroxysteroid dehydrogenases, such as HSD17B1, are dysregulated in NAFLD [298], the mechanistic consequences of this are less clear. Decreased oestrogen levels in post-menopausal women are known to increase the risk of developing NAFLD to the same levels as men [319]. It is possible that, *in vivo*, downregulation of HSD17B1 in liver tissue decreases oestrogen synthesis [320] and, in turn, decreases oestrogen signalling, raising the possibility that this enzyme is involved in NAFLD pathogenesis.

Additionally, there was a downregulation of mRNA expression of a number of cytochrome P450 enzymes in response to LPO treatment, including CYP7A1 and CYP3A4. CYP7A1 is involved in the conversion of cholesterol to bile acids and expression is suppressed in NASH [321]. This was found to be a consequence of intestinal fibroblast growth factor 15 (FGF19) activity, leading to farnesoid X receptor (FXR) activation and suppression of CYP7A1 transcription. FGF19 expression increased 1.7-fold in response to LPO treatment, which may explain the suppression of CYP7A1 (**Supplementary file MCS1 – DEGs in LPO-treated HLCs**). It may also indicate that, *in vivo*, liver-based expression of FGF19 plays an important role in suppression of CYP7A1. This may be of importance for NAFLD pathogenesis as CYP7A1 plays a central role in cholesterol metabolism, with deficiencies leading to hypercholesterolaemia [322]. Further to this, there was downregulation of CYP3A4 expression in LPO-treated HLCs, which appears contradictory to the sustained CYP3A4 activity observed in these cells. As CYP3A4 activity is diminished in NASH [323], this this may indicate that LPO-treated HLCs are transitioning from a NAFLD- to NASH-like phenotype, with transcriptional downregulation preceding functional changes. However, the mechanism(s) underpinning the association between hepatic lipid accumulation and suppression of CYP3A4 activity remains unclear.

In human NAFLD, the expression of the UDP-glucuronosyltransferases (UGTs) and sulfotransferases (SULTs) is highly modulated depending upon disease stage [316,324]. UGTs and SULTs represent classes of endoplasmic reticulum-residing drug metabolising enzymes, the majority of which are produced in the liver [316]. There is little evidence to suggest why the genes encoding these enzymes are differentially regulated in the presence of lipid accumulation. However, their downregulation, alongside that of many cytochrome P450 enzymes may represent an obstruction, whereby cholesterol cannot be efficiently converted to bile acids or steroid hormones. This, in turn, could increase availability of intracellular cholesterol, making it available for incorporation into lipid droplets [325]. Furthermore, there was a decrease in expression of 5β-reductase (AKR1D1), which in human NAFLD is progressively downregulated with advancing disease, leading to intracellular lipid accumulation [315]. Together with decreased expression of cytochrome P450 enzymes, UGTs and SULTs, these changes in mRNA expression may, in part, explain the development of macrovesicular steatosis in LPO-treated HLCs.

In addition to these observed changes, there was a profound decrease in the expression of a number of histone structural units, including H1 and H2A. Histone depletion, such as this, has been observed extensively in yeast, where it either leads to transcriptional repression or de-repression [326]. More recently, in human cells, histone depletion was found to impact mRNA expression and cell cycle progression [327]. This raises the question of whether histone depletion initiates the widescale changes in transcription observed here of whether it is a consequence of the NAFLD phenotype. If it is the former, this further raises the question of what causes histone depletion. While it is not understood why NAFLD would impact on expression of histone structural units, other studies have observed similar perturbations. A variant of histone H2A is macroH2A1.1, which was previously found to protect against lipid accumulation when overexpressed in HepG2 cells, suggesting that depletion may be associated with lipid accumulation [328]. Similarly, induction of lipid accumulation in human hepatoblastoma cells results in depletion of histone structural unit H1 [42], indicating an interplay between intracellular lipid and histone structural units. This study also hypothesised that loss of histone structural units makes cells more

vulnerable to DNA damage from excessive reactive oxygen species (ROS) formation [42]. Whilst ROS was not measured directly in the study presented here, transcriptomic data revealed an upregulation of endoplasmic reticulum oxidoreductase 1β (EROB1) (**Supplementary file MCS1 – DEGs in LPO-treated HLCs**), which is regulated by the ER unfolded protein response (UPR) [329] and increases ROS generation in NAFLD [330–332]. This suggests that steatotic HLCs may increase ROS generation, which may lead to increased DNA damage. Further study within this area may be informative for understanding the causes of DNA mutations giving rise to NAFLD-associated HCC.

Another source of excessive ROS formation in cells is respiratory complex dysfunction in the mitochondrial ETC [333], with complexes I, II and III producing significant levels when dysfunctional [334]. There is evidence from both mouse [335,336] and human [337] studies of defects in the ETC in NAFLD and NASH, leading to increased ROS production. In line with these observations, there was extensive dysregulation of genes associated with oxidative phosphorylation and the TCA cycle in steatotic HLCs. Previous mouse or indirect human studies described increases in hepatic TCA cycle activity associated with NAFLD [34,109,129]. Additionally, compromised ETC activity, with decreased mitochondrial maximal respiration, was previously observed in response to hepatic steatosis [130]. Significant perturbation of genes within these pathways suggest that steatosis may be associated with altered TCA cycle activity, and this will be explored further in the following chapter.

3.5 Experimental limitations

Although HLCs are morphologically and functionally similar to hepatocytes, they do not necessarily recapitulate the entire transcriptome of human hepatocytes [338]. Whilst LPO-treated HLCs express a substantial number of NAFLD-related genes, it is unclear how expression patterns are impacted by the lack of co-culture with other liver-resident cells, such as stellate or Kupffer cells. Development of co-culture systems, or 3D models, using pluripotent stem cell-derived stellate or Kupffer cells,

will facilitate further understanding of how cell-to-cell communication impacts the transcriptional programme of steatotic HLCs. Likewise, as with other *in vitro* models, HLCs are not exposed to the endocrine cues that they would normally receive *in vivo*, particularly those that are generated as a result of obesity.

3.6 Conclusion

Taken together, the data presented in this chapter demonstrate that LPO-treated HLCs not only accumulate intracellular lipid but do so in a pattern similar to that observed in human pathology, demonstrating the utility of this model in the study of NAFLD. Furthermore, transcriptional analyses reveal similarities with human disease, and also reveal extensive dysregulation in pathways related to mitochondrial energy metabolism. In light of these findings, I chose to use this model to further analyse the impact of NAFLD on TCA cycle metabolism and downstream effects on TET enzyme activity.

Chapter 4: Understanding the link between lipid accumulation and mitochondrial metabolism in steatotic HLCs

4.1 Introduction

NAFLD has been linked to impaired mitochondrial respiration, with previous findings revealing altered electron transport chain (ETC) activity associated with steatosis [130]. Multiple studies suggest that the pathology of NAFLD is linked to impairment of mitochondrial function [34,109–111]. In rat models, mitochondrial dysfunction precedes the development of hepatic steatosis, suggesting that disease pathology and progression are linked to impairment of mitochondrial function [32]. Further, studies in mice indicate that hepatic steatosis is accompanied by increased tricarboxylic acid (TCA) cycle activity [109,129]. In support of this, limiting pyruvate transport into the mitochondria reduces mitochondrial flux and limits the development of NASH-associated inflammation [339]. Human studies also suggest that NAFLD is associated with increased TCA cycle activity, with increased movement of substrates into and out of the TCA cycle alongside increased rates of gluconeogenesis and lipolysis [34]. However, whilst informative, these human studies were indirect and used measurements of plasma metabolites to infer hepatic metabolic flux, rather than directly measuring levels within the liver.

The use of whole tissue for metabolic flux analysis is challenging and confounded by the presence of heterogeneous cell populations, each of which may have a different metabolic phenotype [340]. Studies in primary rat hepatocytes show that modulation of TCA cycle anaplerosis fuels oxidative stress and enhances development of a NASH phenotype [299]. One limitation of primary cells is that they rapidly lose their phenotype and display altered transcriptional signatures on purification [300,301]. In these studies, we employed a renewable form of human liver tissue [341]. To my knowledge, high-resolution metabolomic analyses have not been performed in a human system, and the data presented here will provide the field with new insight into human liver disease.

4.1.1 Hypothesis

Nutrient excess in NAFLD leads to mitochondrial dysfunction and subsequent increases in TCA cycle activity due to increased availability of TCA cycle substrates, including pyruvate.

4.1.2 Aims

1. To determine whether macrovesicular steatosis is associated with mitochondrial metabolic dysfunction
2. To assess TCA cycle activity in response to macrovesicular steatosis
3. To examine the interplay between perturbed TCA cycle activity and macrovesicular steatosis

4.2 Materials and methods

4.2.1 Differentiation of hPSCs to HLCs and induction of lipid accumulation

Human female H9 pluripotent stem cells (hPSCs) were differentiated to hepatocyte-like cells (HLCs), and lipid accumulation induced as described in Chapter 3. For isotopic tracing studies, lactate was replaced with $^{13}C_3$-lactate (CK Isotopes, CLM-1579-05). For mechanistic studies, HLCs were exposed to either 5-Aminoimidazole-4-carboxamide-1-β-D-ribofuranosyl 5′-monophosphate (AICAR; 1mM; Sigma-Aldrich, A1393-50MG), O-(Carboxymethyl)hydroxylamine hemihydrochloride (AOA; 100 μM ; Sigma-Aldrich, C13408-1G) or AICAR combined with monomethyl fumarate (50 μM; Sigma-Aldrich, 651419-1G) for the same duration as LPO.

4.2.2 Mitochondrial stress test

The oxygen consumption rate (OCR) of LPO-exposed HLCs was measured using the Agilent Seahorse XF Cell Mito Stress Test Kit (Agilent, 103015-100) on a Seahorse XF Analyser (Agilent, California, USA). The Mito Stress Test provides a readout of oxygen consumption in cells and uses this as a proxy for assessing mitochondrial ETC activity. Analysis was performed under basal conditions and following treatment with oligomycin A (an ATPase inhibitor), carbonyl-cyanide-4-(trifluoromethoxy) phenylhydrazone (FCCP; an ETC uncoupler), and combined rotenone and antimycin A (inhibitors of complex I and III, respectively). Inhibition of ATPase by oligomycin prevents the flow of protons from the mitochondrial intermembrane space and subsequent generation of ATP, revealing levels of ATP-linked respiration occurring within the cell. Addition of FCCP uncouples the ETC, allowing the free flow of protons through the inner mitochondrial membrane, helping to determine the maximum respiratory capacity of the mitochondria. Finally, inhibition of complexes I and III with combined rotenone and antimycin A is used to reveal the level of non-mitochondrial respiration occurring within the cell. Two concentrations of FCCP (0.5 μM and 1.0 μM) were used for optimisation. Since replicates within each group responded similarly to each other, results were combined. OCR was normalised to total protein for each well, using the sulforhodamine B (SRB) assay, as described in Chapter 2.

4.2.3 Gas chromatography-mass spectrometry and nuclear magnetic resonance spectroscopy

These protocols were performed by the labs of collaborators, Dr Daniel Tennant and Dr Christian Ludwig, in the Institute of Metabolism and Systems Research at the University of Birmingham. Experimental details are outlined in Chapter 2.

4.2.4 Transcriptomic analysis

RNA was purified from HLCs that were incubated with either LPO or LPO + AICAR. This was sent to the Wellcome Trust Clinical Research Facility (Edinburgh, UK), where 3' RNA sequencing was performed. Differential expression and principal component analyses were then performed using the Bluebee platform from Illumina (Netherlands). Heatmaps were generated with Heatmapper [287]. Pathway enrichment analysis was performed using DAVID [291,292].

4.2.5 Statistics

All statistical analyses were performed using Graph Prism Version 8.0 for Windows or macOS, GraphPad Software, La Jolla California USA, www.graphpad.com. Normality of data distribution was measured using the Shapiro-Wilks test. Where indicated, data were analysed by unpaired Student's t-test, Mann-Whitney test or two-way analysis of variance (ANOVA). Data were considered to be significant where $p < 0.05$.

4.3 Results

4.3.1 Macrovesicular steatosis in HLCs is associated with electron transport chain dysfunction

In the previous chapter, I showed that macrovesicular steatosis in HLCs was linked to disruption of genes associated with the TCA cycle and oxidative phosphorylation. This raised the question of whether this translated into a functional impact on mitochondrial respiration. In order to test this further, I performed a mitochondrial stress test, using a Seahorse Analyser (**Figure 4.1**).

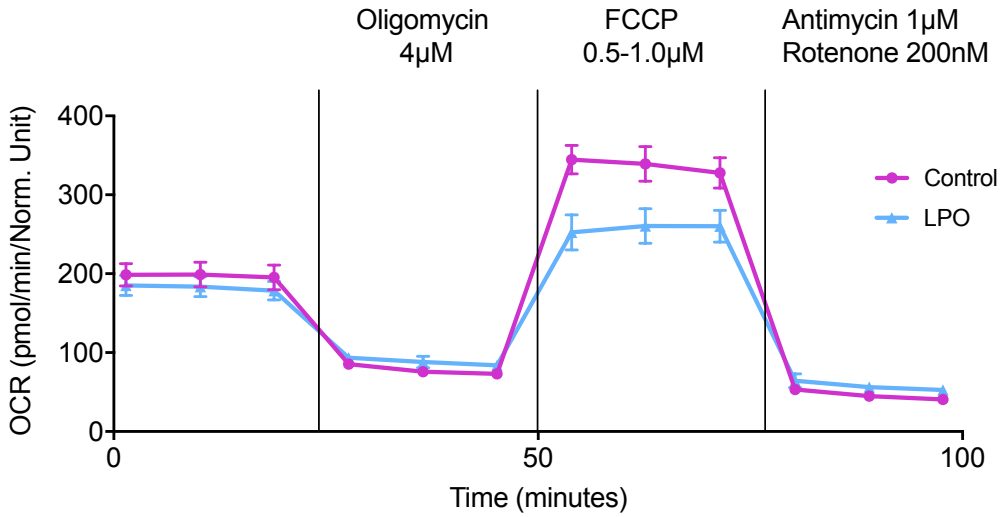


Figure 4.1. Raw trace of oxygen consumption rate (OCR) of control or LPO-treated HLCs.

The basal oxygen consumption rate (OCR), representing combined mitochondrial and non-mitochondrial oxygen consumption, was unchanged following LPO exposure (**Figure 4.2A**). Additionally, oligomycin A, a complex V (ATPase) inhibitor decreased OCR equally effectively in both the control and treatment group. Firstly, this indicates that there were no changes in ATP-linked respiration in response to macrovesicular steatosis (**Figure 4.2B**). Secondly, when comparing oligomycin A-induced alterations in OCR with those following addition of rotenone/antimycin A, I detected no changes in proton leak between groups (**Figure 4.2C**). The addition of the ETC uncoupler FCCP revealed a decrease in maximal respiration in the steatotic HLCs (LPO treated group), suggesting ETC dysfunction (**Figure 4.2D**). Subsequently, complex I and III

were targeted with rotenone/antimycin A, respectively, to completely inhibit oxidative phosphorylation, which reduced OCR to a similar level in both groups, suggesting no difference in non-mitochondrial sources of OCR between the control and steatotic HLCs. Using OCR measurements following FCCP treatment and comparing to the basal OCR, I calculated that there was a decrease in reserve capacity in the LPO-treated cells, compared with control cells (**Figure 4.2E**). Data adapted from Sinton *et al* (2020) [303].

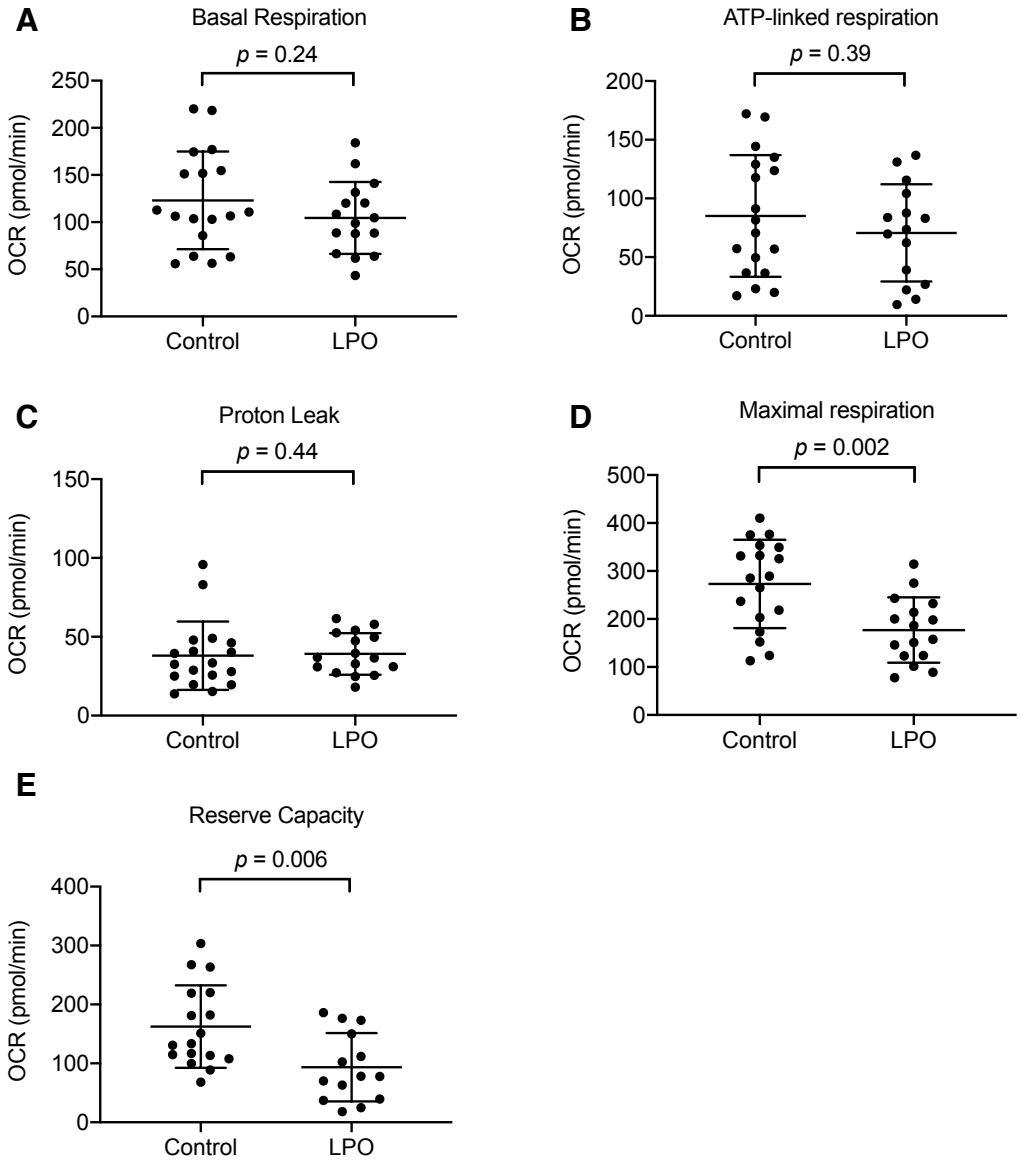


Figure 4.2. LPO treatment results in decreased maximal respiration in mitochondria but not in other aspects of oxygen consumption rate. HLCs were injected sequentially from ports A-C with 4 μM oligomycin, 0.5-1 μM FCCP, and 1 μM antimycin combined with 200 nM rotenone. (**A**) basal respiration; (**B**) ATP-linked respiration; (**C**) proton leak; (**D**) maximal respiration; (**E**) reserve capacity compared to control treatment (n = 18 and 16 biological replicates in the control and treatment group, respectively). Data analysed using two-tailed Student t-test for parametric data, or Mann-Whitney U test for non-parametric data, and expressed as mean ± SD. Data adapted from Sinton *et al* (2020) [303].

4.3.2 Induction of macrovesicular steatosis is not associated with loss of mitochondrial integrity or alterations in mitochondrial number

To analyse whether the observed changes in OCR could be due to loss of mitochondrial membrane integrity rather than ETC dysfunction, I measured citrate synthase (CS) activity. LPO treatment appears to have no impact on CS activity (**Figure 4.3**), suggesting that mitochondria remain intact following intracellular lipid accumulation. Although the sample size here is low, with a large spread in the data, it correlates with my other findings, suggesting that this is accurate. However, repetition of this assay, with a greater sample size, would be useful. I next measured abundance of specific mitochondrial proteins, succinate dehydrogenase subunit A (SDHA) and pyruvate dehydrogenase (PDH) α1 and α2 subunits, and observed no change following treatment with LPO (**Figure 4.4A-B**).

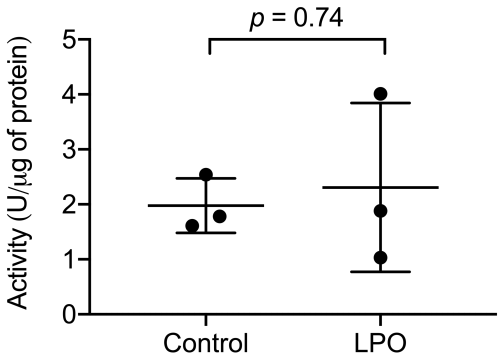


Figure 4.3. LPO does treatment does not alter citrate synthase activity. Levels of citrate synthase activity were unaltered by LPO treatment. Data analysed using two-tailed Student t-test and expressed as mean ± SD, n = 3 biological replicates/group. Data adapted from Sinton *et al* (2020) [303].

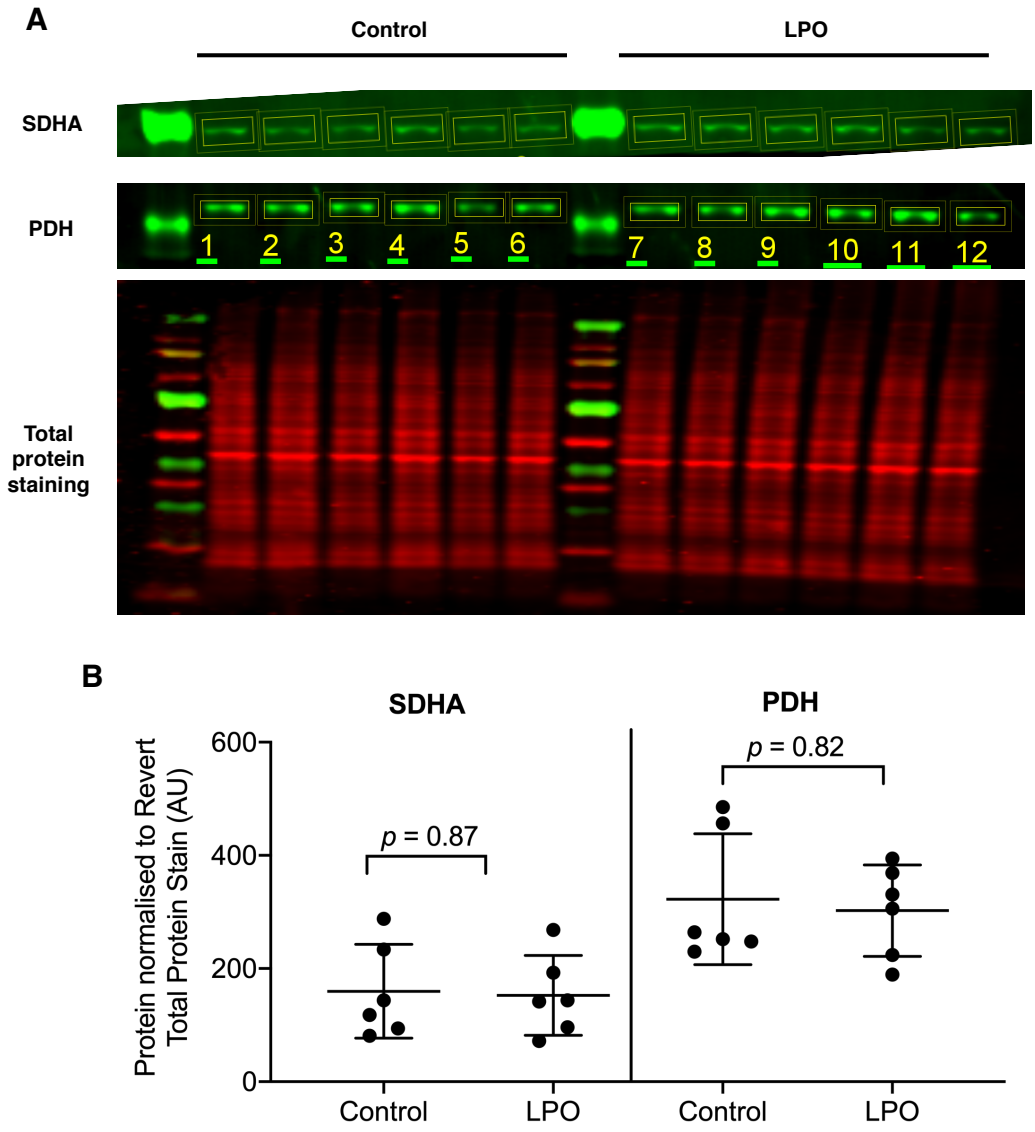


Figure 4.4 LPO treatment does not alter levels of specific mitochondrial proteins. There were no changes in abundance of the mitochondrial proteins succinate dehydrogenase subunit (SDHA) or pyruvate dehydrogenase α1 and α2 subunits (PDH). (**A**) Representative blot. (**B**) Protein quantification. SDHA = succinate dehydrogenase subunit A; PDH = pyruvate dehydrogenase. Data were analysed using a two-tailed Student t-test and expressed as mean ± SD. Data adapted from Sinton *et al* (2020) [303].

A final question was whether changes in ETC function, as demonstrated by decreased cellular maximal respiration and reserve capacity, were due to decreased numbers of mitochondria in response to intracellular lipid accumulation. Since these were the only respiratory measurements to change, I hypothesised that this was not the case. In support of this, there were no changes in mitochondrial or nuclear DNA (**Figure 4.5A**) and no change in the ratio of nuclear to mitochondrial DNA (**Figure 4.5B**). Finally, high content microscopy was used to measure intracellular mitochondrial quantity, by measuring MitoTracker fluorescence, and revealed no changes following exposure to LPO (**Figure 4.6A-B**). Taken together, these data suggest that mitochondrial number is not altered in the presence of intracellular lipid accumulation.

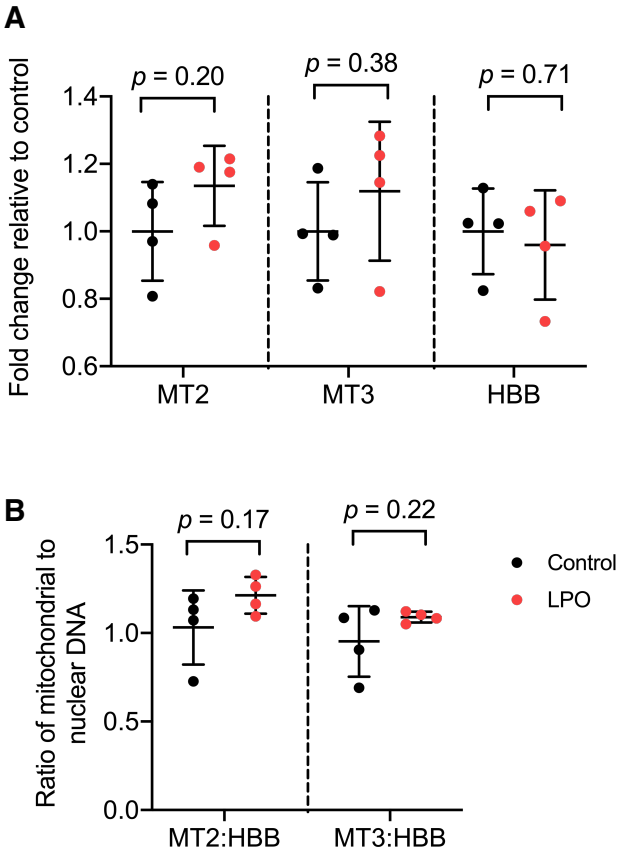


Figure 4.5. LPO treatment does not alter ratio of mitochondrial to nuclear DNA. No changes were observed in abundance of DNA in (**A**) mitochondrial region 2 (MT2), mitochondrial region 3 (MT3) or genomic region beta globulin (HBB). There were no significant changes in the ratio of mitochondrial to genomic DNA (**B**), indicating no alterations in the quantity of mitochondria in response to treatment. Data analysed using two-tailed Student t-test for parametric data, or Mann-Whitney U test for non-parametric data, and expressed as mean ± SD, n = 4 biological replicates/group. Data adapted from Sinton *et al* (2020) [303].

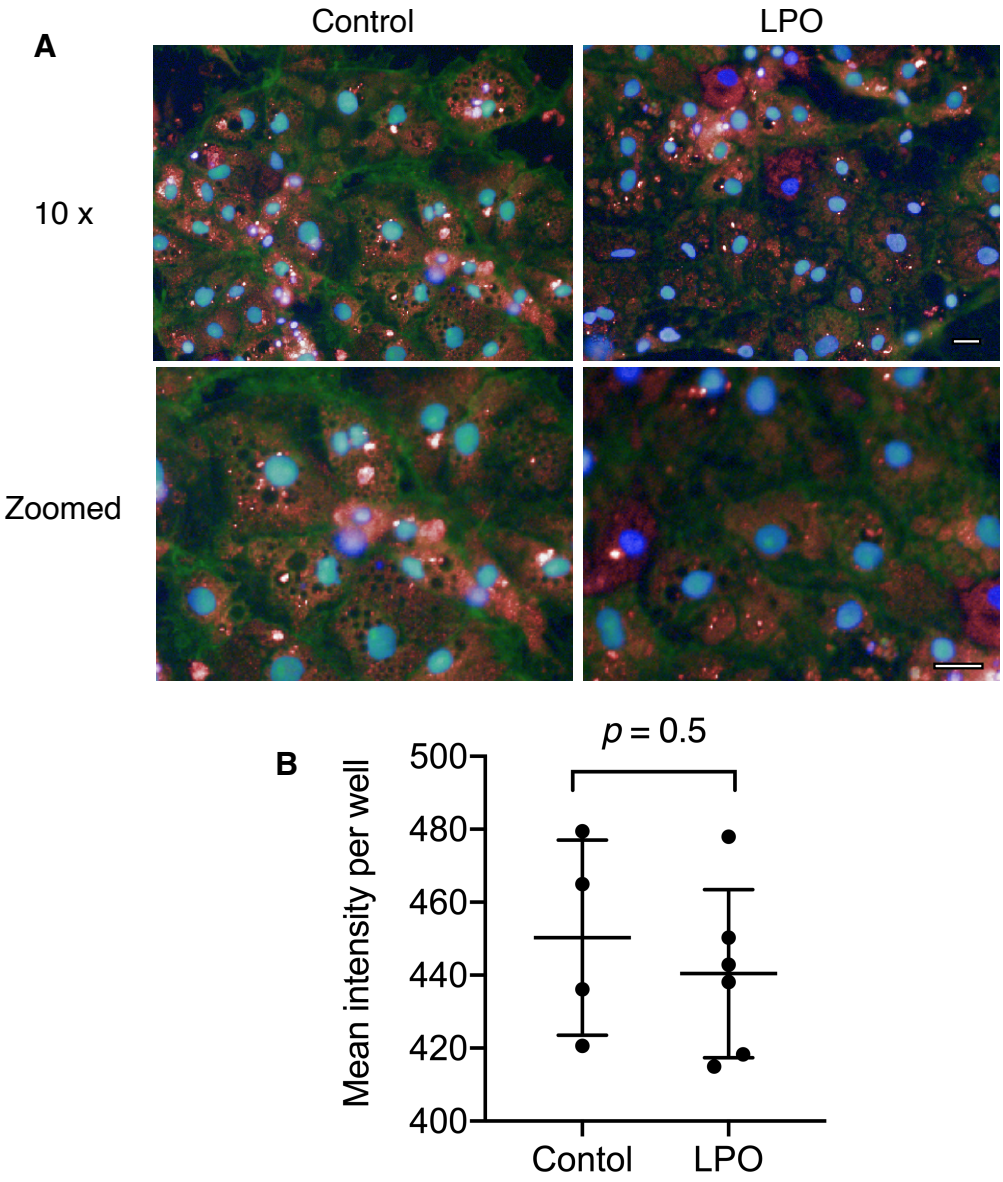


Figure 4.6. LPO treatment does not induce changes in the quantity of mitochondria. (**A**) Representative images of mitochondrial content in control & LPO-treated groups 10x magnification or zoomed in on the same images. Blue staining = NucBlue, red staining = MitoTracker Deep Red; (**B**) high content microscopy revealed no changes in fluorescence in response to treatment (n = 4 and 6 biological replicates in the control and treatment group, respectively). Data analysed using two-tailed Student t-test for parametric data, or Mann-Whitney U test for non-parametric data, and expressed as mean ± SD. Scale bar = 25 μm. Data adapted from Sinton *et al* (2020) [303].

4.3.3 Induction of a NAFLD-like phenotype in HLCs is associated with increased pyruvate carboxylase activity

To investigate TCA cycle dynamics, lactate in the LPO cocktail was replaced with $^{13}C_3$-lactate. Lactate can enter the TCA cycle *via* mitochondrial pyruvate metabolism, through either pyruvate carboxylase (PC) or PDH. Alternatively, pyruvate can be transported into the cytosol and used in gluconeogenesis (**Figure 4.7**).

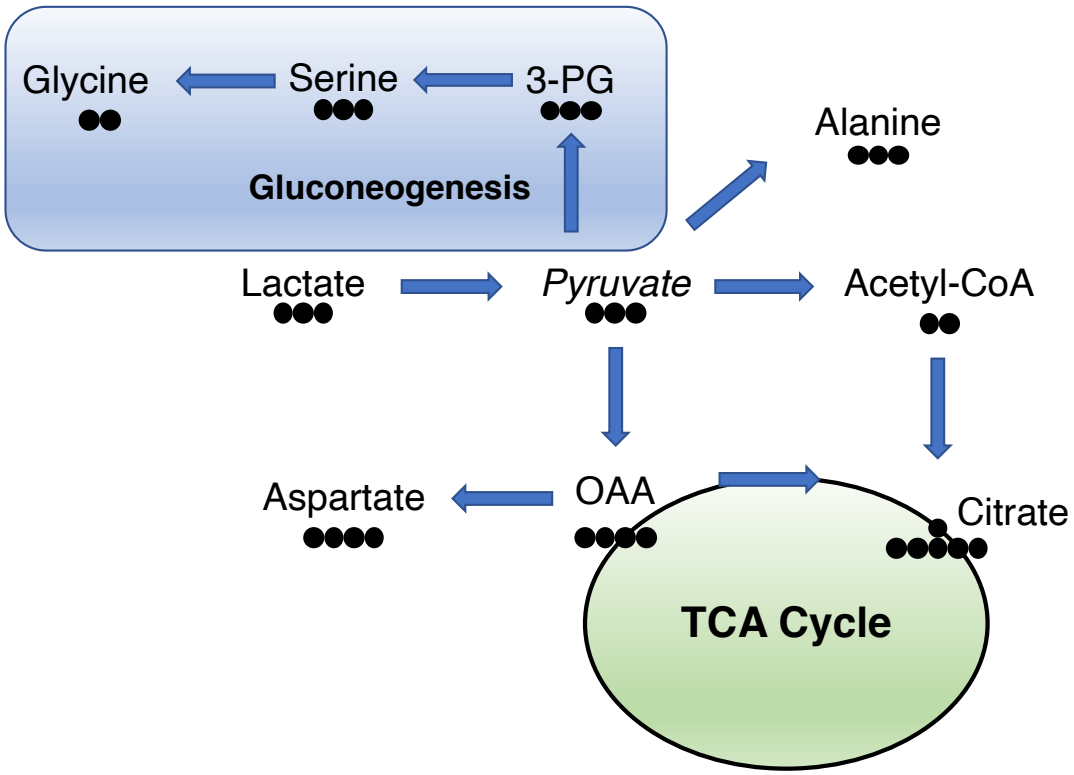


Figure 4.7. Schematic showing the metabolites to which lactate can be converted, following conversion to pyruvate. Black circles indicate the number of carbons in each metabolite. OAA = oxaloacetate.

LPO exposure was associated with increases in steady state levels of pyruvate, aspartate and citrate, but there were no changes in metabolites associated with the gluconeogenesis pathway (**Figure 4.8**). Pyruvate can be converted to alanine, or it can enter gluconeogenesis or the TCA cycle, and so I wanted to determine how it was being used in steatotic HLCs. Whilst there were no changes in the steady state levels

of alanine, there were moderate, but significant, changes in the incorporation of ^{13}C into alanine (**Figure 4.9**), indicating sustained interconversion of alanine and pyruvate.

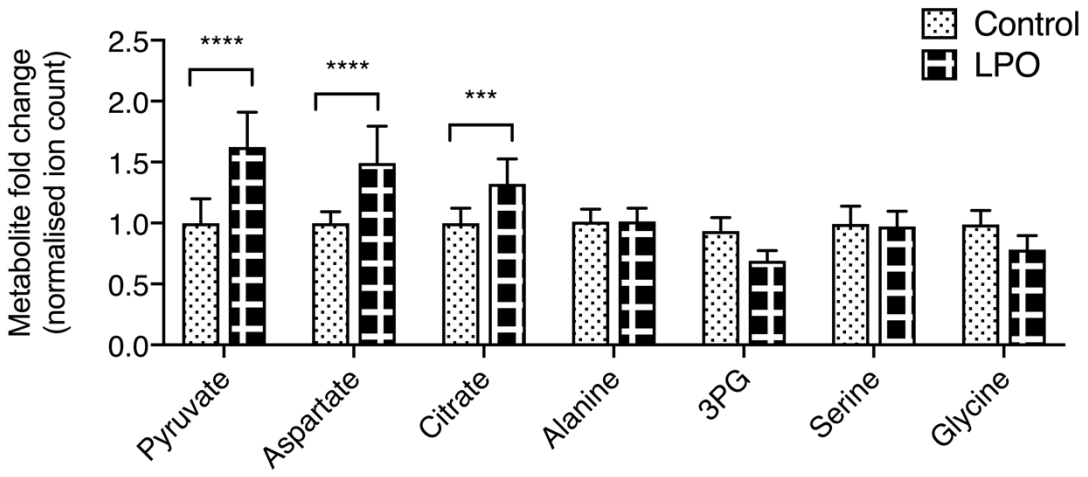


Figure 4.8. Analysis of steady state metabolites shows that in steatotic HLCs, pyruvate is primarily converted to metabolites associated with the TCA cycle, indicating increased PC activity. TCA-associated metabolites: aspartate and citrate. Gluconeogenesis-associated metabolites: 3PG, serine and glycine. Each group consisted of 10 biological replicates within 2 technical replicates (n = 10 replicates/group). Data were analysed by two-way ANOVA with Sidak post-hoc testing and are expressed as mean ± SD, $^{**}p<0.01$, $^{****}p<0.0001$. Figure adapted from Sinton *et al* (2020) [304].

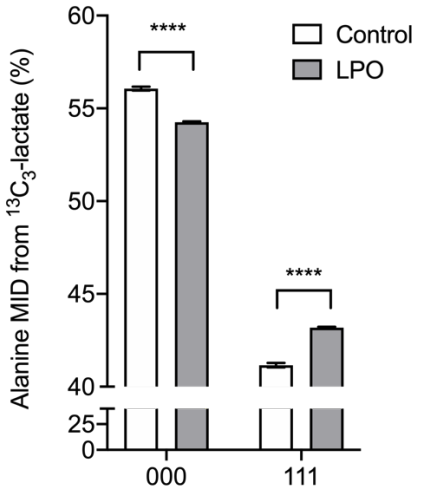


Figure 4.9. NMR spectroscopy analysis reveals minimal conversion of pyruvate to alanine in steatotic HLCs. MID (multiple ion detection mode) enables determination of proportions of different isotopomers in a pool. Alanine is a 3-carbon molecule; here each 0 represents an unlabelled carbon and each 1 represents a ^{13}C-labelled carbon. *n* = 4 biological replicates/group. Data were analysed by two-way ANOVA with Sidak post-hoc testing and are expressed as mean ± SD, ****$p<0.0001$. Figure adapted from Sinton *et al* (2020) [304].

Despite there being no increase in steady state levels of gluconeogenesis-associated metabolites, the increased m+2 labelling of 3-PG (**Figure 4.10A**), serine (**Figure 4.10B**), and glycine (**Figure 4.10C**), suggests increased flux through the gluconeogenesis pathway with LPO exposure. Although pyruvate was added as part of the LPO cocktail, increased pyruvate generation also occurred as a result of increased lactate dehydrogenase activity in LPO exposed cells as shown by increased incorporation of ^{13}C into the metabolite (**Figure 11A**). It was not possible to directly measure conversion of pyruvate to OAA, so I used aspartate as a surrogate and found increased m+3 labelling in steatotic cells. This indicates direct conversion of pyruvate to OAA and, therefore, increased PC activity with LPO exposure (**Figure 4.11B**). Furthermore, isotopomer labelling of citrate indicated preferential synthesis from OAA and decreased conversion of labelled acetyl-CoA by PDH (**Figure 4.11C**).

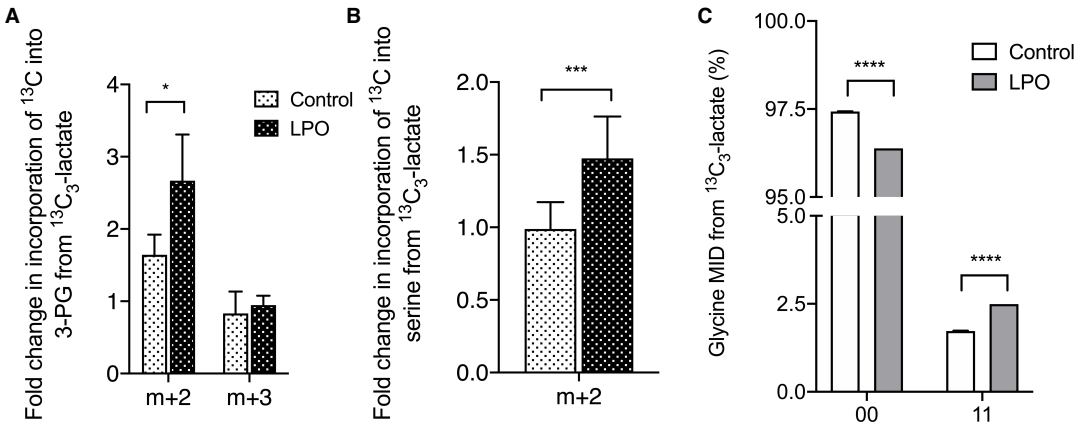


Figure 4.10. Increased flux of m+2 isotopomers indicates sustained conversion of pyruvate. In steatotic HLCs there was increased incorporation of ^{13}C into m+2 isotopomers of (**A**) 3-PG, (**B**) serine, and (**C**) glycine. GC-MS data (**A-B**) consisted of 10 biological replicates within 2 technical replicates (n = 10 replicates/group). For NMR data (**C**), each 0 represents an unlabelled carbon and each 1 represents a ^{13}C-labelled carbon; n = 4 biological replicates/group. Data were analysed by two-way ANOVA with Sidak post-hoc testing, and are expressed as mean ± SD, *p<0.05, ***p<0.001, ****p<0.0001. Figure adapted from Sinton *et al* (2020) [304].

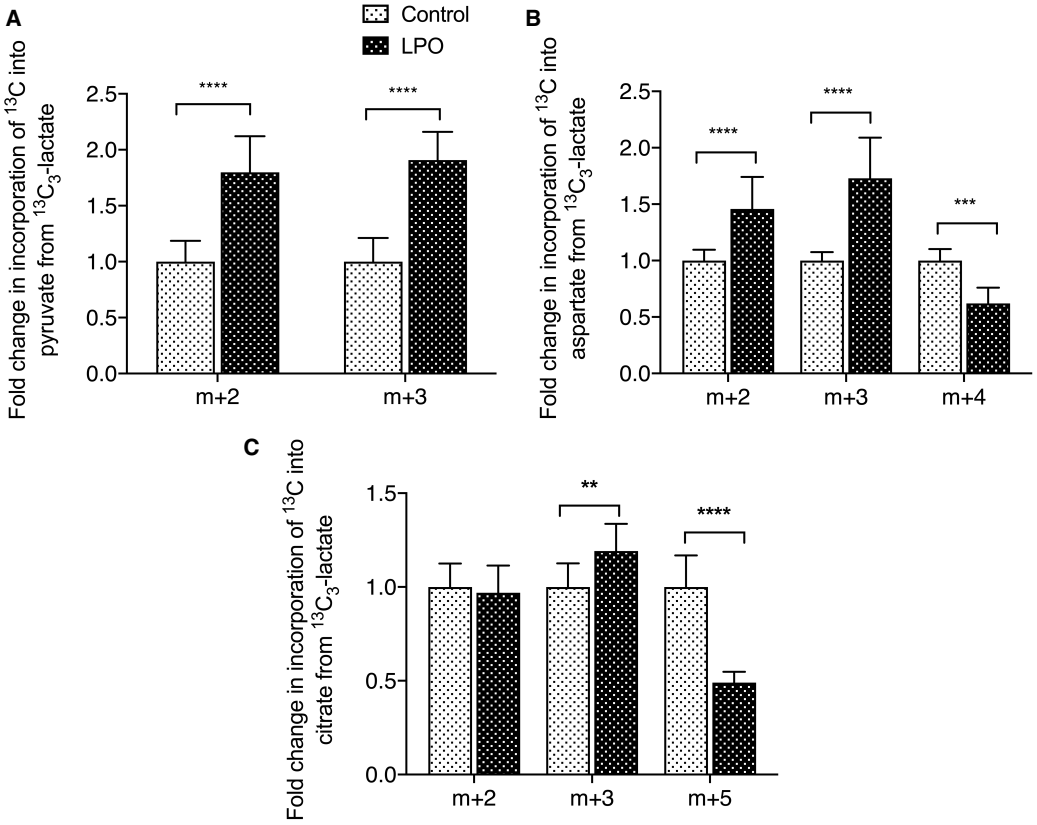


Figure 4.11. Steatotic HLCs display increased pyruvate carboxylase activity. Isotopomer labelling patterns of pyruvate (**A**), aspartate (**B**) and citrate (**C**) reveal that $^{13}C_3$-lactate is preferentially utilised by the TCA cycle. Data consisted of 10 biological replicates within 2 technical replicates (n = 10 replicates/group). Data were analysed by two-way ANOVA with Sidak post-hoc testing, and are expressed as mean ± SD, $^{**}p<0.01$, $^{***}p<0.001$, $^{****}p<0.0001$. Figure adapted from Sinton *et al* (2020) [304].

4.3.4 Steatosis in HLCs is associated with truncation of the TCA cycle

Isotopic labelling of TCA metabolites can produce a number of different isotopomers, dependent on the directionality of metabolite synthesis (**Figure 4.12**). In steatotic HLCs, there was an increase in the steady state levels of akg, fumarate and malate, but no changes in succinate or glutamate (**Figure 4.13**). The lack of changes in steady state levels of glutamate in response to LPO indicate that there is no increase in cataplerosis through this metabolite. This raised the question of why succinate levels

do not change and indicated the presence of more complex changes occurring in the TCA cycle than were previously appreciated.

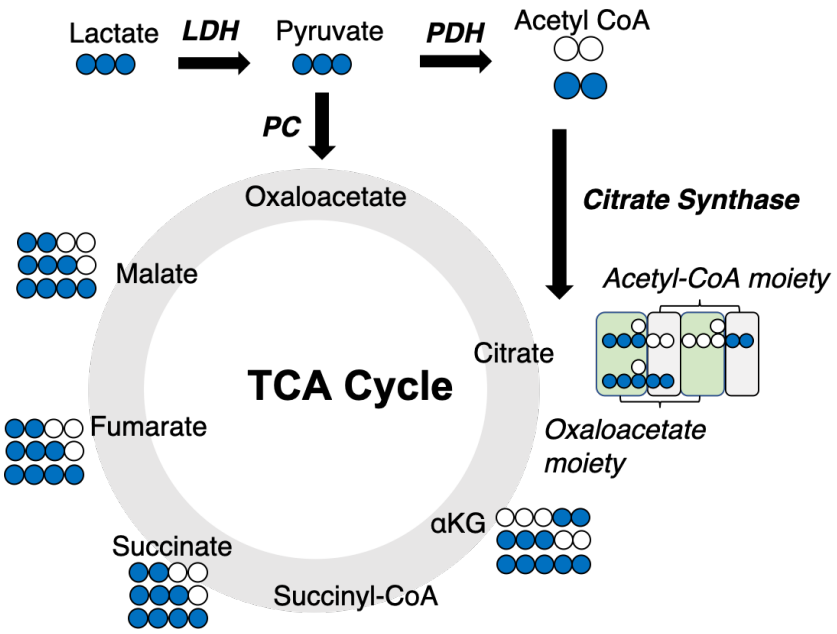


Figure 4.12. Schematic outlining the TCA cycle and possible ^{13}C labelled isotopomers from a single cycle. Blue circles = ^{13}C; white circles = ^{12}C. Figure adapted from Sinton *et al* (2020) [304].

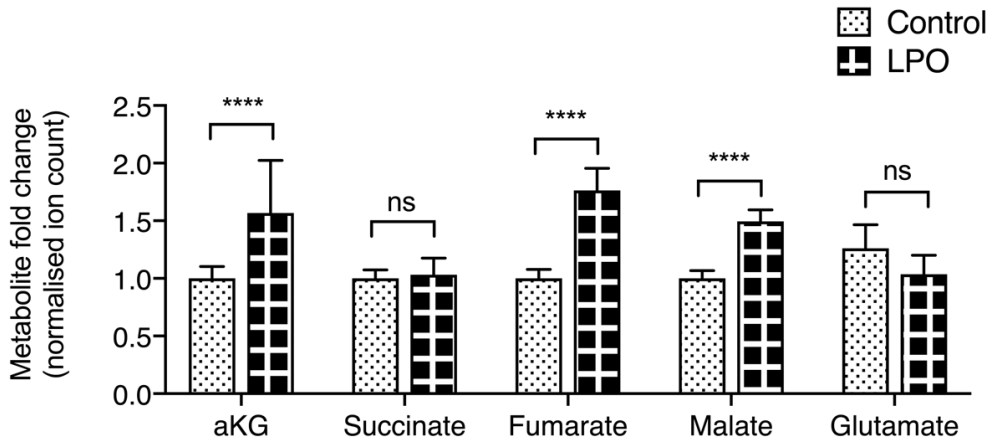


Figure 4.13. Analysis of steady state metabolites by GC-MS reveals increased abundance of αKG, fumarate and malate, but not succinate. All Control and LPO group data consisted of 10 biological replicates within 2 technical replicates (n = 10 replicates/group). Data were analysed by two-way ANOVA with Sidak post-hoc testing and are expressed as mean ± SD, ****p<0.0001. Figure adapted from Sinton *et al* (2020) [304].

To further interrogate these changes, ^{13}C incorporation into TCA cycle-associated metabolites was measured. Whilst there was an increase of ^{13}C incorporation into PC-derived citrate (**Figure 4.11C**), this was not the case for αKG (**Figure 4.14A**). I then examined whether this was due to increased cataplerosis through glutamate. However, there were decreased levels of ^{13}C incorporation into glutamate demonstrating that this is not the case (**Figure 4.14B**). Increased expression of OGDHL and decreased expression of SUCG1, as shown in the Chapter 3, Figure 3.15, suggest that steatosis is associated with the generation of an increased pool of succinyl-CoA, with impaired conversion to succinate. This is supported by substantially decreased incorporation of ^{13}C into succinate (**Figure 4.14C**). In contrast, there was increased ^{13}C incorporation into fumarate and malate (**Figure 4.15A-B**). As these metabolites can be replenished by the purine nucleotide cycle (PNC) or the malate-aspartate shuttle (MAS), I next wanted to investigate whether either of these pathways was contributing to their generation in steatotic HLCs.

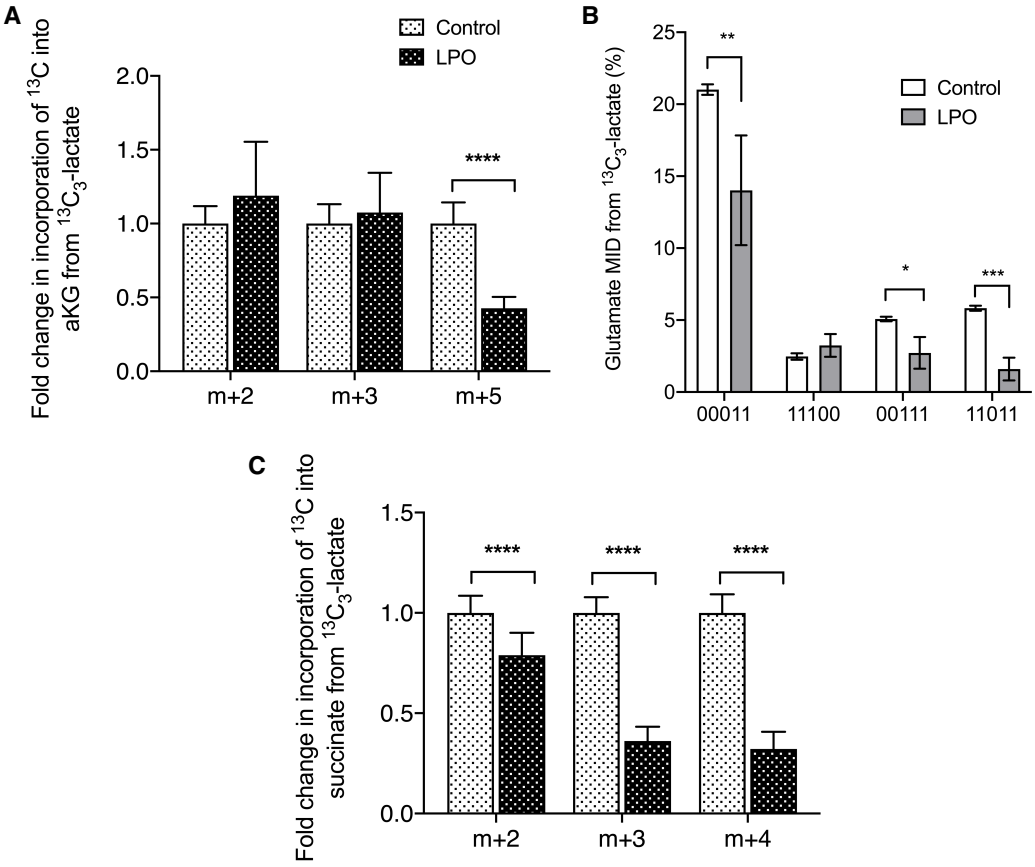


Figure 4.14. PC-derived αKG is unchanged in the presence of steatosis, and PDH-derived αKG is reduced. (**A**) Labelling patterns of αKG are not due to increased cataplerosis through glutamate, as both PC- and PDH-derived glutamate are reduced in response to steatosis (**B**); each 0 represents an unlabelled carbon and each 1 represents a ^{13}C-labelled carbon (**C**) Reduced ^{13}C labelling of succinate indicates inefficient conversion from αKG. (**A, C**) consisted of 10 biological replicates within 2 technical replicates (n = 10 replicates/group). (**B**) n = 4 biological replicates/group. Data were analysed by two-way ANOVA with Sidak post-hoc testing, and are expressed as mean ± SD, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. Figure adapted from Sinton *et al* (2020) [304].

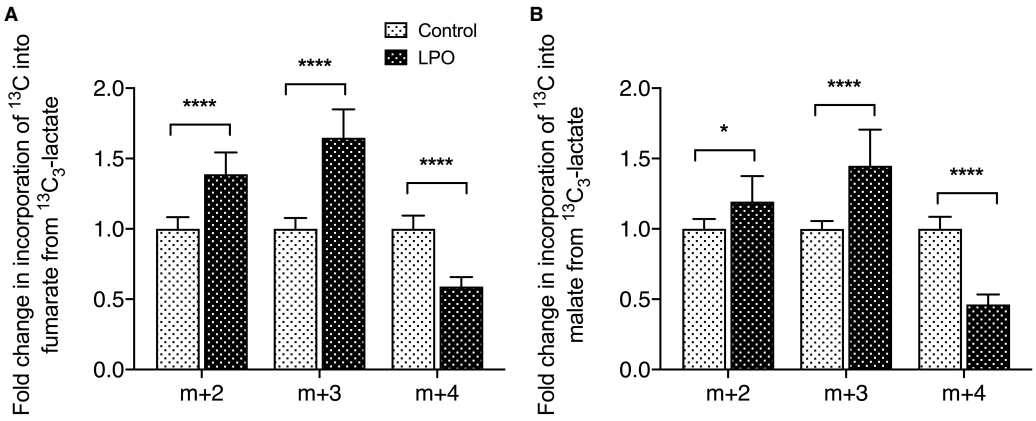


Figure 4.15. In steatotic HLCs, despite TCA cycle truncation, fumarate and malate continue to accumulate. Both fumarate (**A**) and malate (**B**) show increased PC-derived label incorporation. Each group consisted of 10 biological replicates within 2 technical replicates (n = 10 replicates/group). Data were analysed by two-way ANOVA with Sidak post-hoc testing and are expressed as mean ± SD, *p<0.05, ****p<0.0001. Figure adapted from Sinton *et al* (2020) [304].

4.3.5 Increased PNC and MAS activity drives NAFLD-associated fumarate accumulation

Next, I wanted to determine the source of fumarate accumulation in steatotic HLCs. Transcriptomic analysis in the previous chapter identified dysregulation of multiple genes associated with the MAS (**Figure 4.16A**) and PNC (**Figure 4.16B**), indicating that these pathways may be involved in fumarate accumulation. To investigate this further, HLCs incubated with $^{13}C_3$-lactate-labelled LPO were co-incubated with either AICAR or AOA, to inhibit the PNC or MAS, respectively. Addition of AICAR increased the steady state levels of αKG, and neither AICAR nor AOA had a significant impact on the steady state levels of pyruvate, citrate, succinate or malate (**Figure 4.17**). However, both inhibitors significantly reduced steady state levels of fumarate and aspartate.

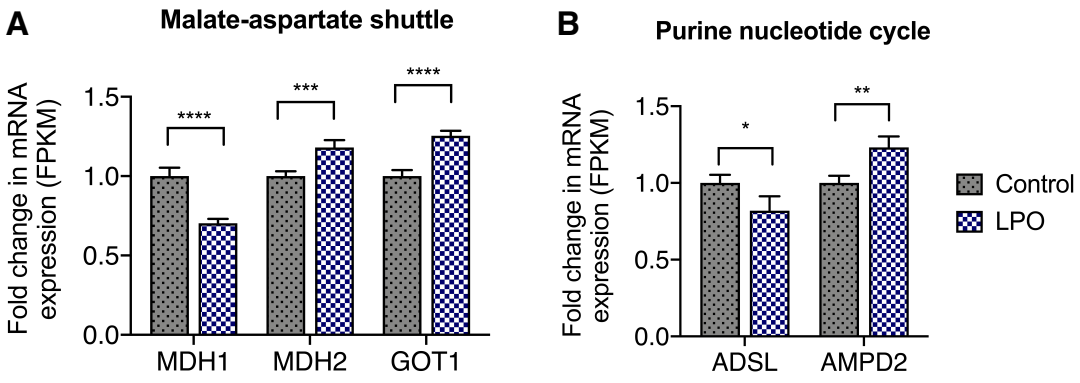


Figure 4.16. Transcriptomic analysis revealed disrupted expression of genes associated with the malate-aspartate shuttle and purine nucleotide cycle. Sample size *n* = 3/group. Data were analysed by Student's t-test and are expressed as mean ± SD. **p*<0.05, ***p*<0.01, ****p*<0.001, *****p*<0.0001. MDH1 = Malate Dehydrogenase cytosolic isoform; MDH2 = Malate Dehydrogenase mitochondrial isoform; GOT1 = Glutamic-Oxaloacetic Transaminase 1; ADSL = Adenylosuccinate Lyase; AMPD2 = Adenosine Monophosphate Deaminase-2 Figure adapted from Sinton *et al* (2020) [304].

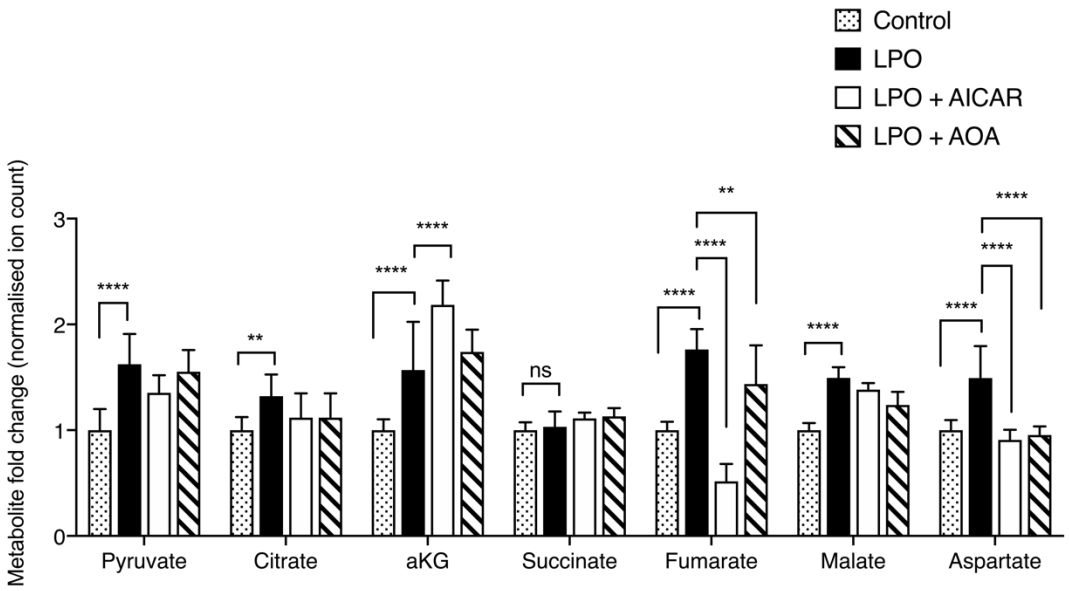


Figure 4.17. Inhibition of the PNC and MAS reversed steatosis-induced accumulation of fumarate and aspartate, but not malate. All Control and LPO group data consisted of 10 biological replicates within 2 technical replicates (*n* = 10 replicates/group), as shown in previous GC-MS figures. GC-MS LPO + AICAR and LPO + AOA groups consisted of 6 biological replicates/group. Data were analysed by two-way ANOVA with Sidak post-hoc testing and are expressed as mean ± SD, ***p*<0.01, ****p*<0.001, *****p*<0.0001. Figure adapted from Sinton *et al* (2020) [304].

Whilst addition of AICAR did not impact the synthesis of PDH-derived citrate (**Figure 4.18A**), it did result in increased accumulation of PDH-derived ^{13}C in αKG (**Figure 4.18B**), suggesting impaired conversion to succinyl-CoA or increased anaplerosis from glutamate. AICAR reversed LPO-induced cataplerosis of αKG through glutamate (**Figure 4.18C**) and was also able to partially restore steatosis-associated depletion of succinate (**Figure 4.18D**). This resulted in small but significant increases in PC- and PDH-derived succinate isotopomers (m+4 and m+2, respectively). The most profound effect of AICAR was on fumarate, with each isotopomer reduced to levels below those observed in the control group (**Figure 4.19A**), demonstrating that in steatotic HLCs fumarate accumulation is primarily driven by the PNC. Following AOA treatment, there was a moderate decrease in PC-derived ^{13}C incorporation into

fumarate, demonstrating a small contribution from the MAS. AICAR had a moderate impact on incorporation of PC-derived incorporation of ^{13}C into malate, whereas AOA impacted on both PC- and PDH-derived ^{13}C incorporation, suggesting limited contribution of the MAS and PNC to malate accumulation (**Figure 4.19B**). Both AICAR and AOA reduced incorporation of ^{13}C into aspartate (**Figure 4.19C**).

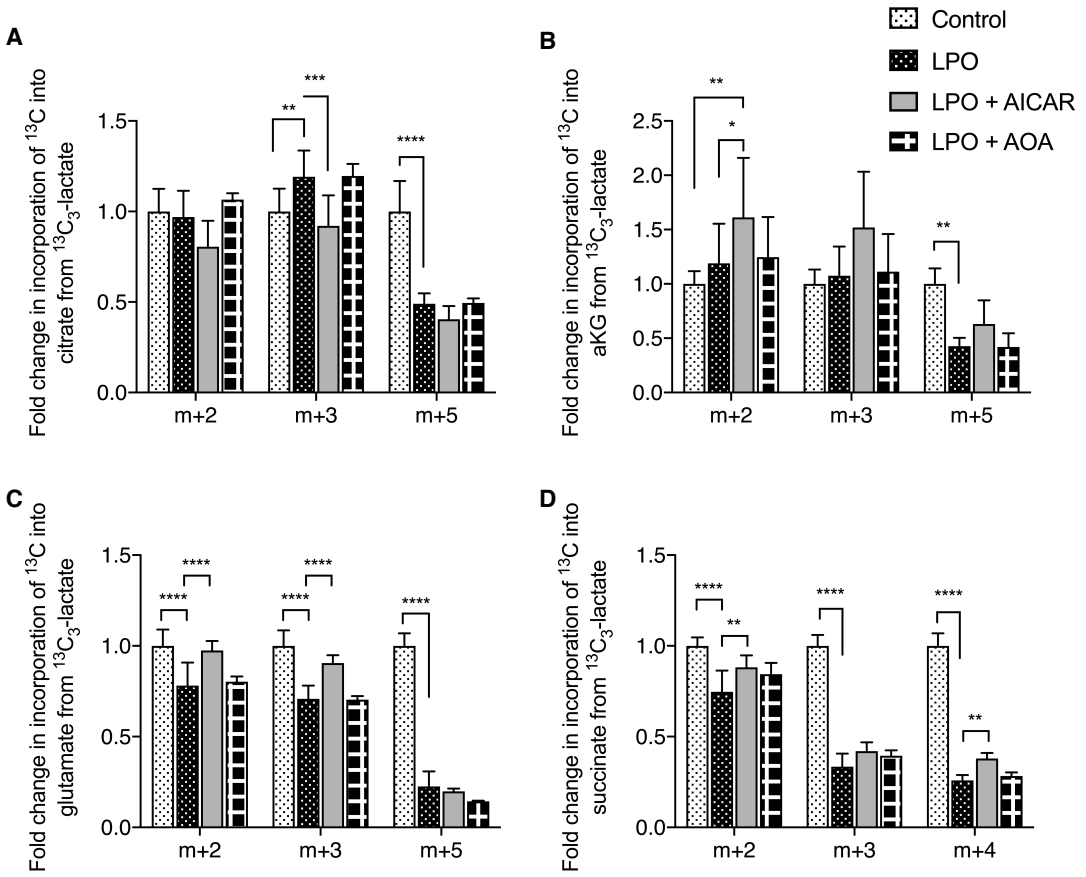


Figure 4.18. PNC inhibition restores glutamate cataplerosis but does not reverse TCA cycle truncation. PNC inhibition moderately reduced PC-derived citrate ^{13}C labelling (**A**), but not (**B**) αKG. Despite this, there was a moderate increase in glutamate upon PNC inhibition (**C**). PNC inhibition did not impact on PC-derived succinate levels (**D**). All Control and LPO group data consisted of 10 biological replicates within 2 technical replicates (n = 10 replicates/group), as shown in previous GC-MS figures. GC-MS LPO + AICAR and LPO + AOA groups consisted of 6 biological replicates/group. Data were analysed by two-way ANOVA with Sidak post-hoc testing and are expressed as mean ± SD, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. Figure adapted from Sinton *et al* (2020) [304].

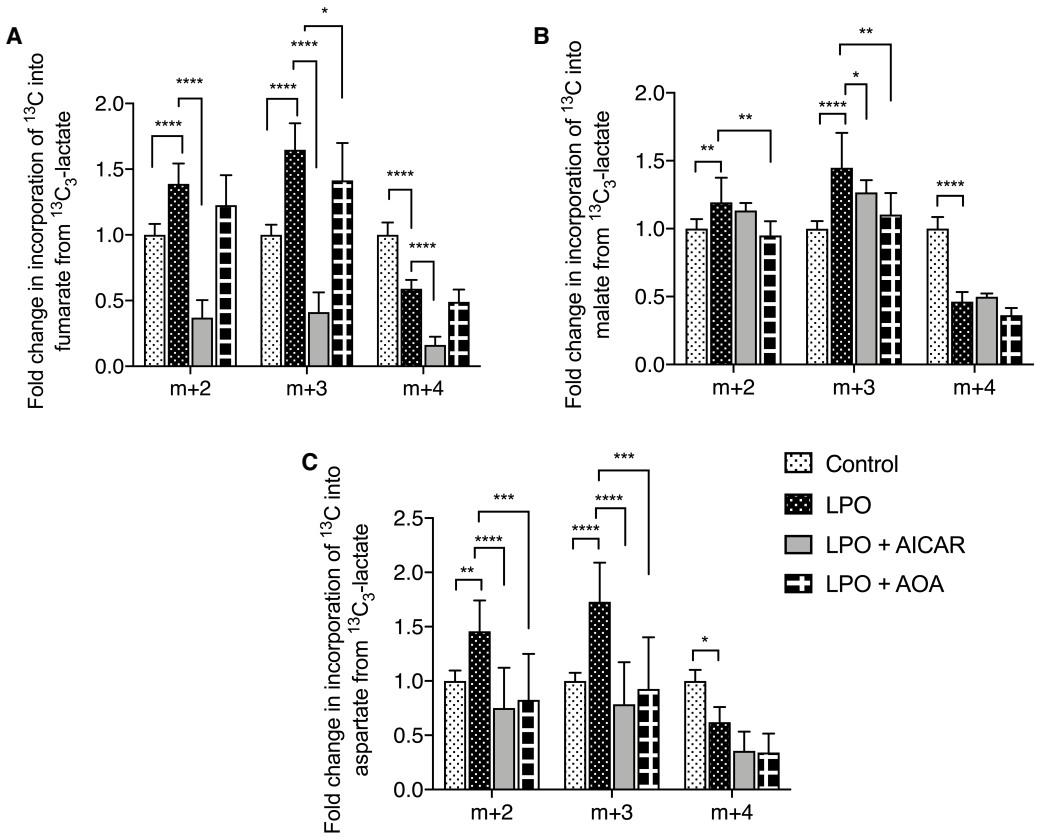


Figure 4.19. The PNC is the major contributor to fumarate in steatotic cells. Inhibition of the PNC, and to a lesser extent the MAS, reduced incorporation of ^{13}C into fumarate in steatotic HLCs (**A**). PNC inhibition also reduced incorporation of ^{13}C into malate and aspartate in steatotic cells (**B-C**). All Control and LPO group data consisted of 10 biological replicates within 2 technical replicates (*n* = 10 replicates/group), as shown in previous GC-MS figures. GC-MS LPO + AICAR and LPO + AOA groups consisted of 6 biological replicates/group. Data were analysed by two-way ANOVA with Sidak post-hoc testing and are expressed as mean ± SD, **p*<0.05, ***p*<0.01, ****p*<0.001, *****p*<0.0001. Figure adapted from Sinton *et al* (2020) [304].

4.3.6 Fumarate accumulation is linked to widescale transcriptional changes in steatotic HLCs

As exposure of HLCs to both LPO and AICAR resulted in reduction of macrovesicular steatosis, I questioned whether this was due to fumarate accumulation influencing transcription. To examine the impact of AICAR on the transcriptome, 3'-RNA sequencing was performed to a depth of >8.8 million reads. Differential expression analysis identified 437 upregulated and 918 downregulated genes, with a Log2 fold change cut-off >0.25 (**Supplementary file MCS2 – DEGs in AICAR-treated HLCs**). Principal component analysis (PCA) demonstrated that replicates within each of the control and LPO group clustered, showing minimal variation between replicates (**Figure 4.20**).

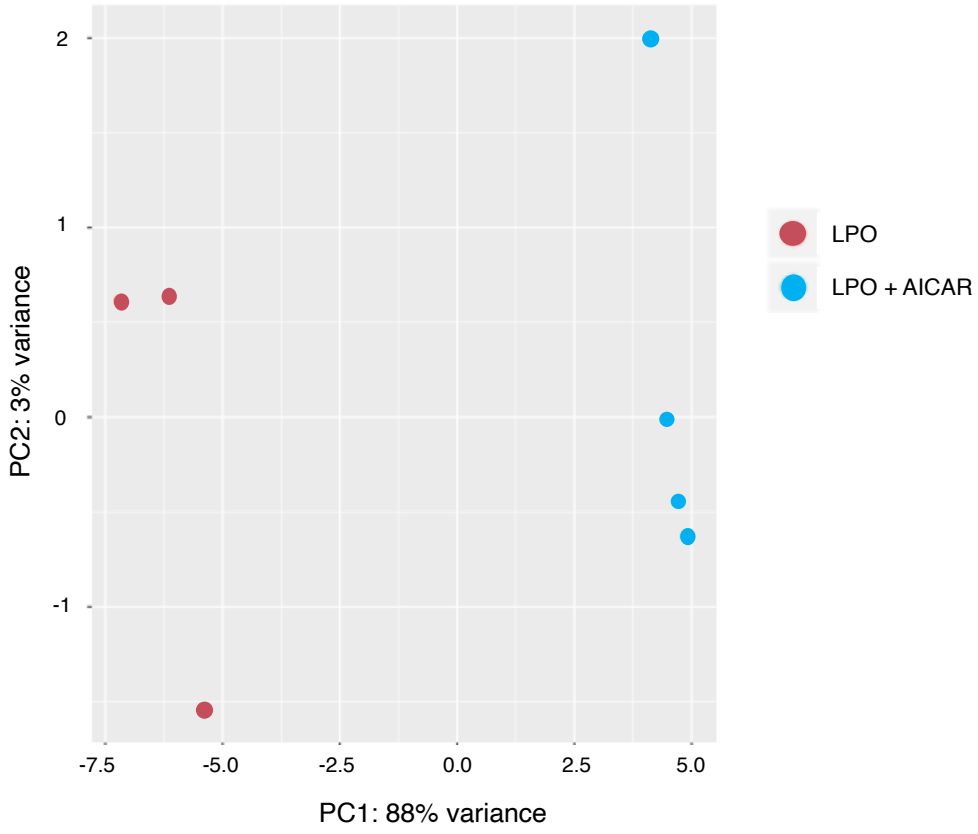


Figure 4.20. Principal component analysis of transcriptomic data reveals clustering of replicates within groups following treatment of HLCs with either LPO or LPO + AICAR. PCA plot demonstrating differential clustering of each biological replicate from both LPO and LPO + AICAR-treated HLCs. For LPO group $n = 3$ biological replicates; for LPO + AICAR group $n = 4$.

Heatmap analysis of dysregulated genes highlights the widescale transcriptomic changes that occur following LPO treatment (**Figure 4.21**). Taken together, these findings demonstrate that LPO + AICAR treatment induces a transcriptional pattern that is distinct from LPO-treated HLCs, and that this is consistent between biological replicates.

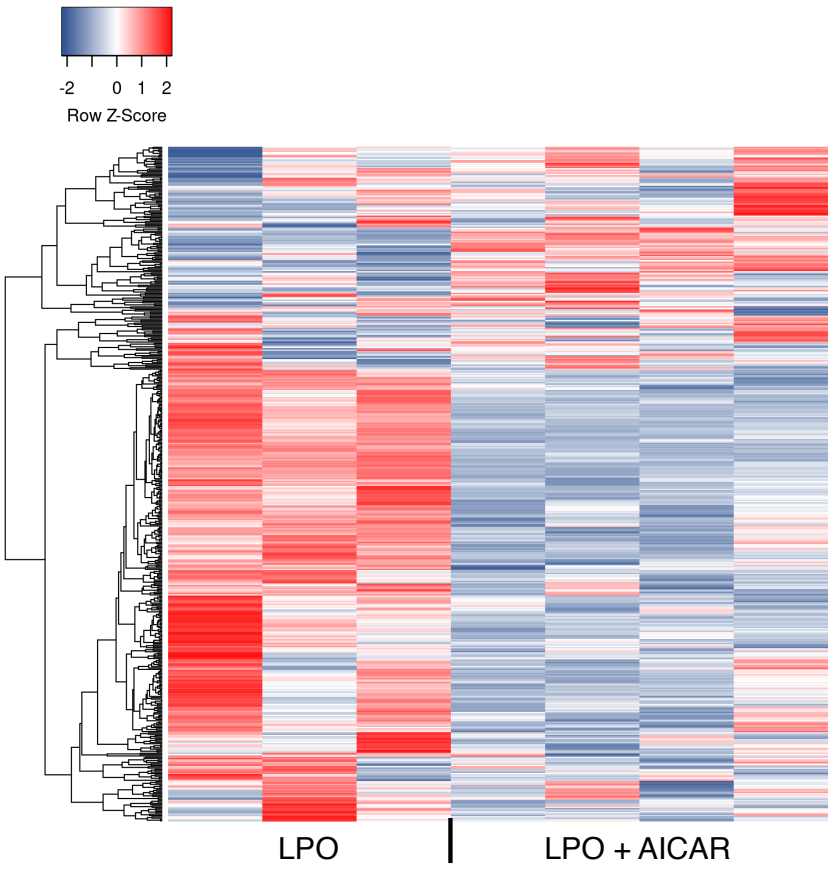


Figure 4.21. Heat map analysis of transcriptomic data reveals wide-spread transcriptomic changes following treatment of HLCs with either LPO or LPO + AICAR. Heap map representing the top 1,000 differentially expressed genes when comparing LPO and LPO + AICAR-treated HLCs. Genes are clustered by Euclidean distance. For LPO group $n = 3$ biological replicates; for LPO + AICAR group $n = 4$.

4.3.7 Reduction of fumarate in steatotic HLCs disrupts pathways related to inflammation and iron ion homeostasis

I mapped genes with a log2 fold change cut-off >0.25 to the Gene Ontology Biological Processes database, identifying a number of enriched pathways. These pathways broadly related to protein translation, immunity and response to oxidative stress (Figure 4.21). These pathways contained genes encoding ribosomal subunits (including RPL7, RPL11, RPS19 and RPS24) components of the complement cascade and acute response to inflammation (including SAA4, ORM1, ORM2, C1R, C3, C5, C8A, C8B and C8G), as well as regulators of iron homeostasis (including HAMP, TFR2, CP, SOD1 and SOD2).

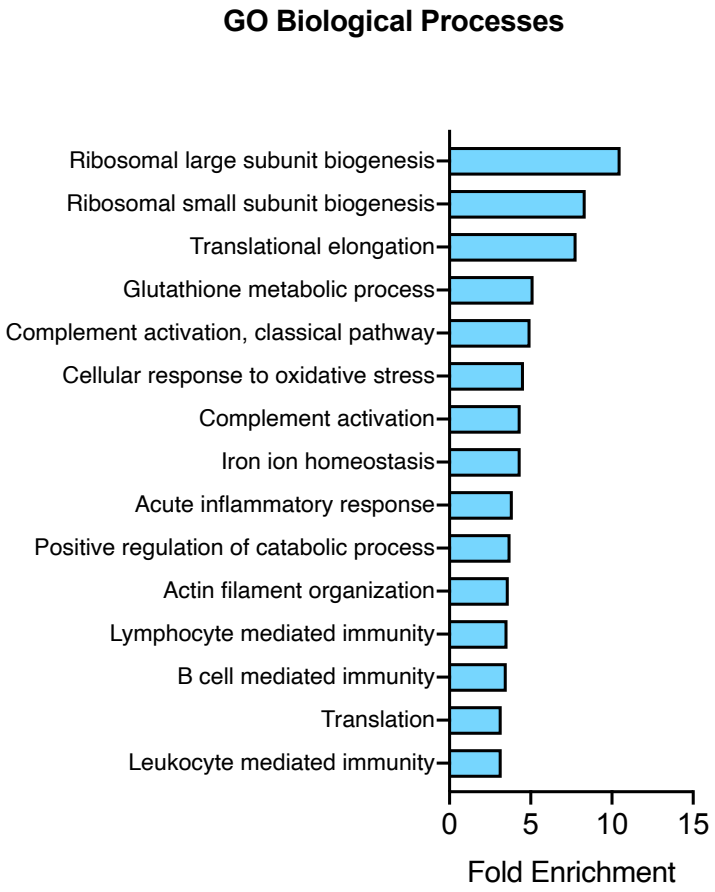


Figure 4.22. Gene ontology (GO) pathway analysis. All genes with a Log_2 fold change > 0.5 were analysed with DAVID. The top 15 pathways for up- and downregulated genes are represented here. Details are located in Appendix 4.

4.3.8 Fumarate accumulation drives macrovesicular steatosis in LPO-treated HLCs

To further examine the relationship between macrovesicular steatosis and PNC-driven fumarate accumulation, I examined transcriptomic data to assess how genes associated with lipid droplet formation respond to concurrent exposure to LPO and AICAR. This revealed a downregulation of the genes PLIN2, PLIN3 and CIDEC (**Figure 4.23**). Further to this, I performed further analyses of lipid droplets using high content analysis microscopy. This revealed that the presence of AICAR alongside LPO was able to prevent the development of macrovesicular steatosis (**Figure 4.24**). Furthermore, the addition of exogenous monomethyl fumarate was sufficient to partially restore the steatotic phenotype, in the presence of both LPO and AICAR. This demonstrates for the first time that fumarate accumulation drives macrovesicular steatosis in LPO-treated HLCs.

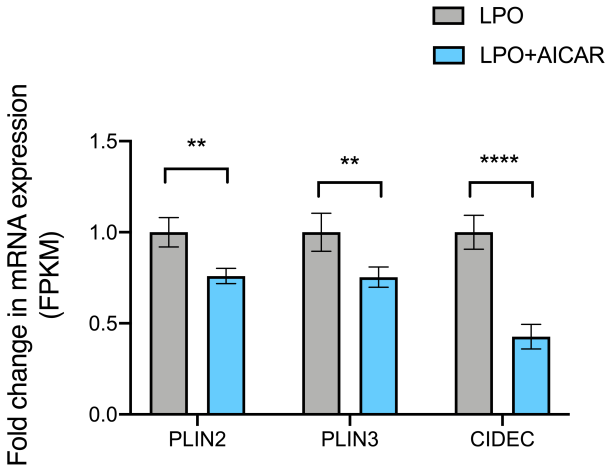


Figure 4.23. Purine nucleotide inhibition is associated with downregulation of genes associated with lipid droplets. Data presented here represent a subset of genes from the RNA-seq dataset (**Supplementary file MCS1 – DEGs in LPO-treated HLCs**). Data were analysed by two-way ANOVA and expressed as mean ± SD. ** $p<0.01$, **** $p<0.0001$

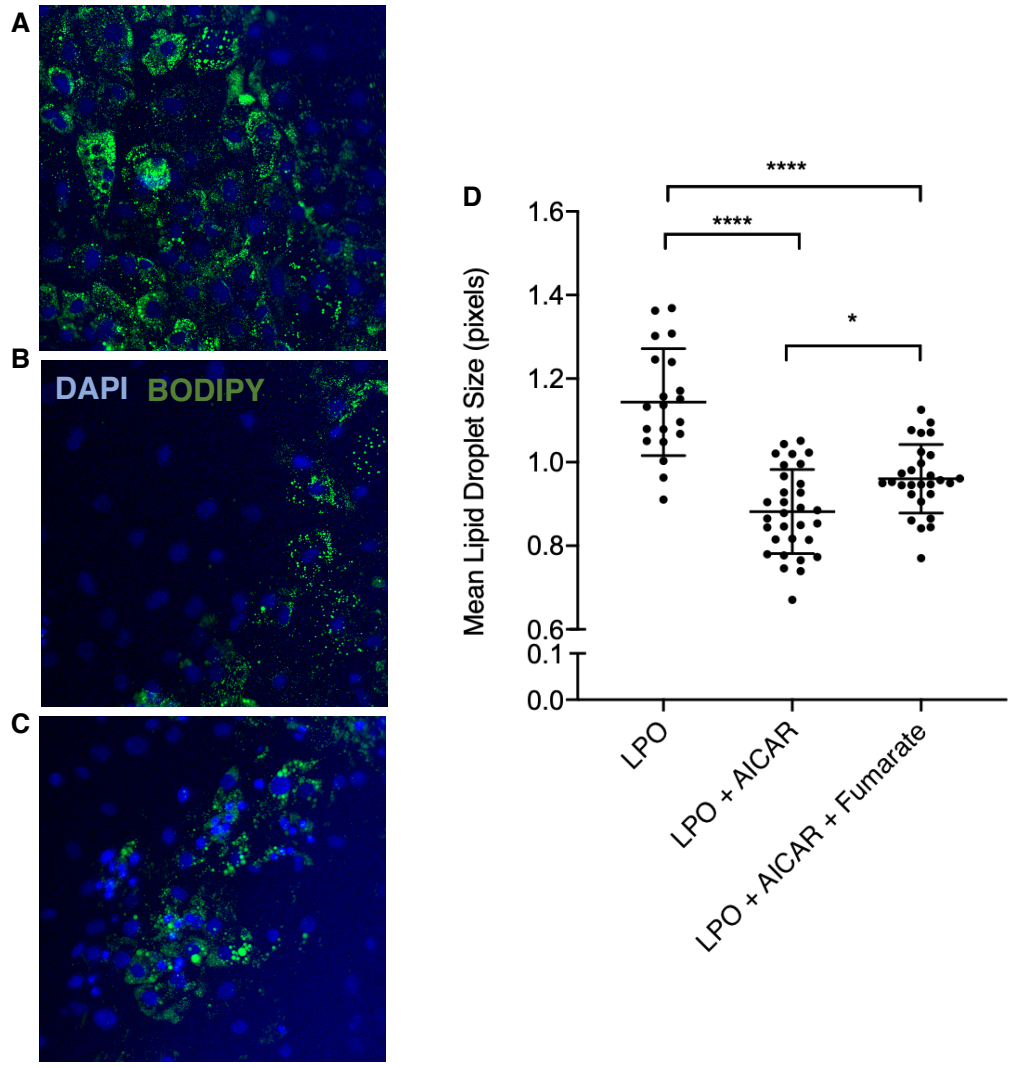


Figure 4.24. Compensatory rewiring of mitochondrial metabolism, resulting in fumarate accumulation, drives development of macrovesicular steatosis in LPO-treated HLCs. (**A**) LPO drives development of macrovesicular lipid droplets; (**B**) Addition of AICAR to LPO-treated HLCs prevents development of macrovesicular lipid droplets; (**C**) Addition of fumarate to HLCs exposed to LPO and AICAR partially restores the macrovesicular steatotic phenotype. (**D**) Quantitative evaluation of lipid droplet size. LPO group $n = 20$, LPO + AICAR group $n = 31$, LPO + AICAR + Fumarate group $n = 28$. Data were analysed by one-way ANOVA with Tukey's multiple comparison test and are expressed as mean ± SD, *$p<0.05$, ****$p<0.0001$. Fumarate = monomethyl fumarate. Figure adapted from Sinton *et al* (2020) [304].

4.4 Discussion

4.4.1 Macrovesicular steatosis is linked to impairment of electron transport chain function

Using an hPSC-based model of NAFLD I was able to overcome many of the difficulties typically associated with studying metabolic dysfunction in this disease. In this chapter I aimed to determine the nature of NAFLD-associated changes in mitochondrial respiration, and whether these occur as a result of mitochondrial dysfunction or reductions in mitochondrial number. Following the establishment of relevance to human disease shown in Chapter 3, I wanted to understand whether macrovesicular steatosis was linked to dysfunctional mitochondrial respiration. Previous studies have shown that mitochondrial maximal respiration increases in obese humans with NAFLD but decreases on progression to NASH, despite an increase in mitochondrial mass [130]. Since NAFLD was previously shown to increase TCA cycle activity [34,109], I hypothesised that mitochondrial respiration would increase in response to LPO treatment. In contrast to this, steatotic HLCs displayed decreased maximal respiration and reserve capacity. This indicates that whilst mitochondria are mostly able to maintain respiratory integrity in LPO-exposed cells, the ETC, and therefore oxidative phosphorylation, may be compromised. Given that these were the only parameters of mitochondrial respiration to change, it suggests that LPO treatment of HLCs inhibits complexes I-III of the ETC. This effect may arise due to the use of octanoic acid, which inhibits complex I-III in rat liver, as well as increasing oxidative stress through reactive oxygen species production [342].

I questioned whether inhibition of complex II (also named succinate dehydrogenase) may impact on TCA cycle activity, due to its role in converting succinyl-CoA to succinate. Furthermore, I questioned whether changes in maximal respiration arose from loss of mitochondrial membrane integrity or resulted from decreased biogenesis. Given that maximal respiration was the only parameter of respiration that changed, I suspected this was not the case. The lack of change in the ratio of nuclear to mitochondrial DNA suggests that mitochondrial biogenesis was not affected by LPO treatment in these studies. This is supported by high content analysis of mitochondria, also showing no changes in mitochondrial quantity, and is consistent with human

studies of NAFLD which show no changes in mitochondrial mass in steatotic versus non-steatotic liver tissue [130]. In contrast, liver tissue from patients with NASH contains greater numbers of mitochondria, with a concomitant decrease in maximal respiration [130]. Therefore, the model of NAFLD presented here may mirror the transition between NAFLD and NASH, with respiratory dysfunction in the absence of mitochondrial biogenesis. This discrepancy may also be due to previous studies using isolated mitochondria from whole tissues, comprised of multiple cell types, whereas we examined a single cell population.

Furthermore, I observed decreased mitochondrial maximal respiration, in the absence of altered mitochondrial biogenesis. Mitochondrial biogenesis is controlled by PPAR-gamma coactivator 1α (PGC1A), with increased expression associated with increased biogenesis [343]. Transcriptomic analysis of steatotic HLCs in Chapter 3 revealed increased expression of PPARGC1A, the gene encoding PGC1A, suggesting that mitochondrial biogenesis should be increased in this context and not unaltered as observed. This may reflect that translation of PGC1A does not strongly correlate to transcription of PPARGC1A. It may also support the assertion that these cells are at a transitional stage, where PPARGC1A expression has increased but has not yet induced mitochondrial biogenesis.

Taken together, these data demonstrate that hepatocyte intracellular lipid accumulation is linked to mitochondrial dysfunction in response to the changing bioenergetic demands associated with NAFLD, impacting on the ETC, and limiting maximal respiration.

4.4.2 Macrovesicular steatosis is associated with increased supply of pyruvate to the TCA cycle

Typically, NAFLD is also thought to increase TCA cycle activity [34,109] but, to date, detailed information regarding flux dynamics has remained elusive. I utilised stable isotopic tracing, using $^{13}C_3$-lactate which provides greater metabolite labelling than

other substrates [344]. In steatotic HLCs, increased lactate dehydrogenase (LDH) activity led to increased generation of pyruvate, a phenomenon which was previously predicted using an *in silico* model of NAFLD [345]. In the liver, flux of pyruvate into the TCA cycle predominantly occurs *via* PC [346], and there is evidence suggesting that PC activity increases in NAFLD [34], contributing to increased rates of gluconeogenesis through conversion of OAA to phosphoenolpyruvate [109]. However, whilst I observed sustained synthesis of gluconeogenesis-associated metabolites occurring in steatotic HLCs, the data presented here show no associated increase in gluconeogenesis. Although increased hepatic gluconeogenesis has been observed in association with NAFLD [347], it appears that it is linked to hepatic IR rather than lipid content [348]. The use of a 2D cell model, and lack of endocrine signalling may prevent IR from developing and explain why I did not observe increased gluconeogenesis in steatotic HLCs. Although studies have stated that NAFLD is the hepatic manifestation of the metabolic syndrome and linked to IR [349], there is evidence that ~50% of patients with this disease do not have IR [350], and the metabolic syndrome can develop in non-obese individuals [351]. Increased incorporation of only two ^{13}C atoms into 3-PG, serine and glycine may suggest that pyruvate-derived acetyl-CoA is being converted back to pyruvate *via* ketogenesis before entering gluconeogenesis. Increased acetyl-CoA pools and disrupted ketogenesis have been observed in NAFLD [347], lending support to this assertion.

Isotopic labelling patterns of aspartate, which is used as a surrogate for OAA, demonstrate that PC activity increases in steatotic HLCs. Although OAA can be converted to phosphoenolpyruvate for use in gluconeogenesis, increased incorporation of $^{13}C_3$ into aspartate and citrate show that pyruvate is being preferentially utilised by the TCA cycle. Since malate dehydrogenase and fumarate hydratase readily reverse their reactions [127,352], increased incorporation of ^{13}C into both malate and fumarate indicate a partial reversal of TCA cycle activity.

4.4.3 Macrovesicular steatosis leads to TCA cycle truncation

The increased flux of pyruvate into the TCA cycle also led to increased TCA cycle activity in steatotic HLCs, as previously reported in humans [34]. However, the data in this chapter show that changes in TCA cycle are more nuanced that was previously suggested. Human studies of NAFLD suggest that the disease is associated with a ~2 fold increase in hepatic mitochondrial oxidative flux, as well as increased levels of gluconeogenesis and lipolysis [34]. However, these claims were made based on indirect studies of plasma from humans with NAFLD. Gluconeogenesis was assessed based on incorporation of deuterium into glucose, and then measuring plasma glucose levels; NAFLD-associated lipolysis was examined by measuring ketone levels in blood plasma. These findings were mirrored in later mouse studies [109,129]. These studies each used [U-^{13}C]propionate or unlabelled propionate to study hepatic TCA cycle flux, due to its conversion to succinyl-CoA and incorporation into the TCA cycle. Importantly, propionate can be converted to pyruvate too, enabling it to enter the TCA cycle via either PC or PDH. Indeed, it was later revealed that administration of propionate increases hepatic pyruvate cycling, TCA cycle activity and gluconeogenesis, throwing into question its suitability as a tracer for studying hepatic metabolism [353]. Furthermore, this study showed that administration of $^{13}C_3$-lactate, in rats, had no impact on plasma levels of succinate, malate or aspartate. It also had no significant impact on gluconeogenesis. Subsequent studies, where rats were fed a high fat diet and developed hepatic steatosis, show similar results when using lactate as a tracer, indicating its utility and that it does not act to drive carbon metabolism independently of other factors [354,355].

Following induction of steatosis in HLCs, similarly to the aforementioned studies of NAFLD, I observed increased steady state levels of the core TCA cycle metabolites, citrate, αKG, fumarate, malate and aspartate. However, steady state levels of succinate were unaltered in steatotic HLCs. This was also reflected in the analysis of isotopomers of these metabolites, raising the question of why less ^{13}C is being incorporated into succinate. Whilst this observation may be due to increased cataplerosis through glutamate, isotopomer analysis revealed decreased flux through this metabolite. Therefore, this indicates that in steatotic HLCs, there may be an

increased pool of succinyl-CoA. Succinyl-CoA could not be measured in this study as CoA profiling requires more sensitive instrumentation, such as liquid chromatography tandem mass spectrometry to be detected [356]. However, transcriptomic data shown in Chapter 3 support this hypothesis. In steatotic cells, OGDHL (2-oxoglutarate dehydrogenase-like), which encodes an enzyme that converts αKG to succinyl-CoA, was upregulated. Further to this, the same analysis revealed that SUCLG1 (succinate-CoA ligase GDP/ADP-forming subunit alpha), which encodes a subunit of an enzyme converting succinyl-CoA to succinate, was downregulated. This suggests that αKG is continuing to be converted to succinyl-CoA in steatotic HLCs, but that disruption of succinate-CoA ligase is preventing further conversion to succinate. Since interconversion between succinyl-CoA and succinate is in near equilibrium and readily reversible [357], I propose that this effect could be a result of diminished succinyl-CoA synthetase activity. In tandem with the findings in section 4.4.1 of this chapter, this reveals a truncation of the TCA cycle in steatotic HLCs, preventing generation of increased succinate levels and preventing efficient conversion of succinate to fumarate. This then raised the question of how increased fumarate levels are being maintained under steatotic conditions.

4.4.4 Mitochondria develop a metabolic shunt to compensate for TCA cycle truncation

In response to truncation, I found that HLCs are able to rewire their metabolic circuitry to compensate for this, generating increased levels of fumarate, predominantly through the PNC, and to a lesser extent the MAS (**Figure 4.24**). A similar metabolic bypass has been reported in cardiac ischemic reperfusion injury [358], as well as human and mouse *in vitro* models of tumorigenesis [127]. However, this phenomenon has not been observed in NAFLD. In ischemic reperfusion injury, increased PNC activity results in fumarate overflow, driving reversal of succinate dehydrogenase activity and accumulation of succinate [358]. I did not observe the same phenomenon here, indicating inhibition of succinate dehydrogenase activity, which is supported by impaired conversion of succinate shown in section 4.4.3 and observed decreases in maximal respiration shown in section 4.4.1. In human and mouse FH-deficient cells,

accumulation of fumarate suppresses complex I activity [127], which may happen in steatotic HLCs too, as indicated by decreased maximal respiration in section 4.4.1.

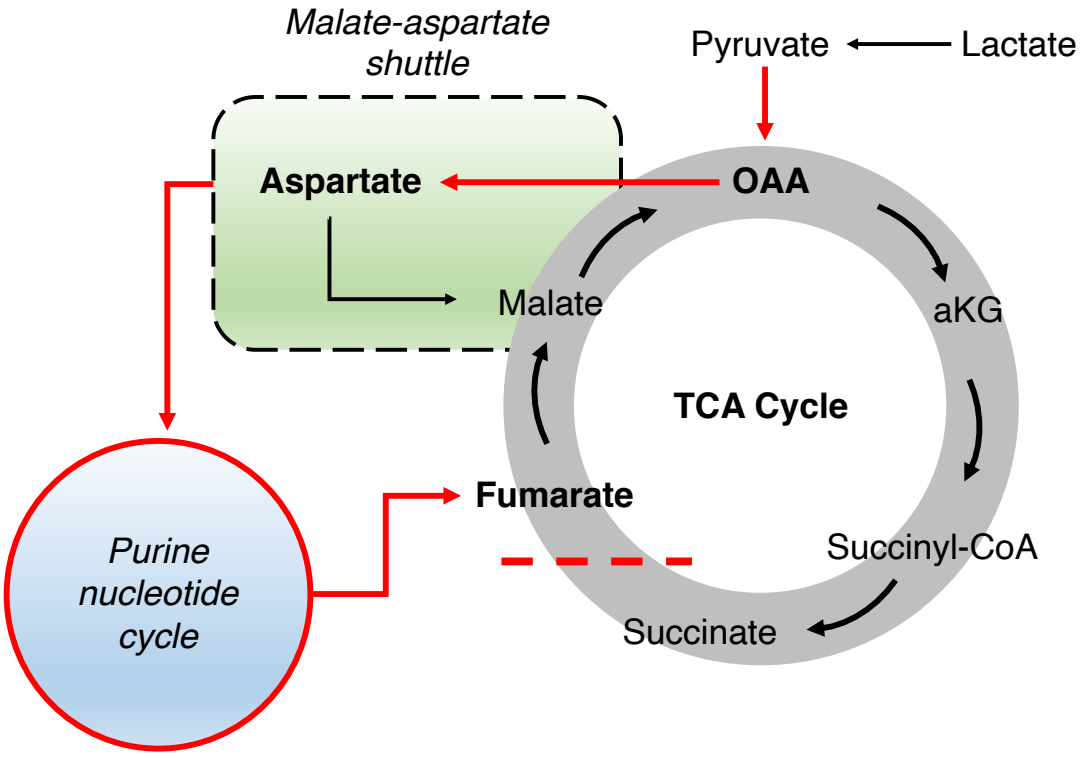


Figure 4.24. Model of metabolic rewiring in mitochondria of steatotic HLCs. Where truncation prevents conversion of succinate to fumarate, the PNC can act as a shunt to provide carbon to the TCA cycle. Solid red lines denote direction of metabolite synthesis in steatotic HLCs. Dashed red line represents TCA cycle truncation. Figure adapted from Sinton *et al* (2020) [304].

Although it is possible for fumarate to be generated from reversal of fumarate hydratase [358], I was unable to directly manipulate the activity of this enzyme in order to assess its contribution to the fumarate pool. However, as FH operates at equilibrium [359] it is likely that low levels of fumarate, resulting from TCA cycle truncation, lead to reversal of enzyme activity.

Defects in complex I can diminish levels of NAD^+ [360] and, therefore, regeneration of fumarate *via* the PNC may be one mechanism of increasing the NADH pool and maintaining the hydrogen ion gradient of the ETC. The synthesis of fumarate by the PNC also results in the generation of AMP, which in turn is deaminated to produce ammonia [361]. In rodents, hyperammonaemia is associated with the progression of hepatic steatosis to cirrhosis [362], and our findings may indicate that this is generated through increased PNC activity. This leads to the question of whether NAFLD-associated cirrhosis is then a consequence of ammonia production or downstream effects of fumarate accumulation. In alcoholic liver disease, administration of monomethyl fumarate reverses hepatic inflammation and steatosis, and has been suggested as a potential therapeutic [363]. Whilst the aetiology of these diseases differs, this is indicative that fumarate plays a role in the progression of NAFLD. To this end, I questioned how inhibition of fumarate accumulation may impact on the development of macrovesicular steatosis.

4.4.5 Inhibition of the PNC is associated with widescale transcriptomic changes

As previously outlined in this chapter, fumarate is a negative regulator of the dioxygenase enzymes, including the TETs and JHDMs. These enzymes are thought to influence transcription through modulation of DNA and histone methylation. Based on this, I hypothesised that fumarate accumulation in steatotic HLCs would have a widescale impact on transcription. Within the transcriptomic dataset from LPO and AICAR-exposed HLCs, I observed changes in the expression of a number of genes associated with complement activation, and regulation of iron homeostasis, most of which were downregulated. It has previously been demonstrated that alterations in complement activation and iron homeostasis pathways are associated with NAFLD [364,365]. Their overall downregulation in response to AICAR treatment may suggest that transcription of genes in these pathways are influenced by epigenetic modifications.

Complement activation is known to be associated with nonalcoholic steatohepatitis (NASH), where complement 3 (C3) accumulates in the liver, where it activates the mannose-binding lectin pathway [364]. This, in turn, recruits neutrophils to the tissue, leading to the development of inflammation [366]. Inhibiting fumarate accumulation in LPO-treated HLCs led to the downregulation of genes associated with the mannose-binding lectin pathway, suggesting prevention of complement activation. This is supported by the finding that DNA methylation of the C3 promoter correlates with transcription [367,368]. In turn, this suggests that fumarate accumulation in NAFLD and the subsequent modulation of dioxygenase enzyme activity can impact on DNA methylation at key inflammation related genes.

The inhibition of fumarate accumulation in steatotic HLCs also influenced expression of genes associated with iron homeostasis. Previous studies suggest that, in humans, hepatic iron accumulation is associated with an increased risk of developing NASH and fibrosis [369][370], but the mechanisms underpinning this association are not yet understood. In NASH patients, serum hepcidin (encoded by HAMP) levels are commonly elevated [371], but it is unclear whether this elevation is a consequence of hepatic disease or associated obesity [372]. In this regard, the transcriptomic data here demonstrate increased expression of HAMP in response to PNC inhibition, indicating that these cells may be releasing iron. Furthermore, transferrin receptor 2 (TFR2) is an iron sensor that can activate hepcidin, and its downregulation in response to PNC inhibition suggests that its expression may be modulated by fumarate in steatotic HLCs. This may indicate that in steatotic HLCs, increased PNC activity promotes iron accumulation. In turn, this may influence transcriptional regulation, as Fe^{2+} is a co-activator of the dioxygenase group of enzymes, including the TETs [137].

4.4.6 Macrovesicular steatosis in HLCs is driven by PNC-induced accumulation of fumarate

A further consequence of fumarate accumulation is the development of macrovesicular steatosis. Inhibition of the PNC prevented development of

macrovesicular lipid droplets, which was partially restored following exposure to exogenous monomethyl fumarate. This correlates with findings in oligodendrocytes and CD8+ T cells, where exposure to exogenous fumarate perturbed lipid metabolism [373,374], although the mechanism by which this occurs remains unknown. However, my data demonstrate that fumarate is important for the development of macrovesicular steatosis in HLCs and that mechanisms compensating for TCA cycle truncation in LPO-treated HLCs cells drive intracellular lipid accumulation.

4.5 Experimental limitations

NAFLD is typically associated with obesity and IR and, therefore, the liver is exposed to a number of adipokines and inflammatory signals *in vivo*. Whilst HLCs develop macrovesicular steatosis, as in other *in vitro* systems they are not exposed to these signals, which may impact their metabolic response. A further limitation is that this study only assessed changes in polar metabolites but did not examine non-polar metabolites, which may give an insight into the response of β-oxidation and ketogenesis pathways to steatosis. This may provide valuable information in understanding how fumarate drives the development of macrovesicular lipid droplets. Furthermore, the development of mass spectrometry imaging could provide insight into the composition of these lipid droplets [375].

A further limitation of the approach taken here is the use of the inhibitor AICAR, which is an adenylosuccinate lyase inhibitor, and blocks PNC activity, thus limiting fumarate accumulation [358]. However, AICAR is also an AMP-activated protein kinase (AMPK) agonist [376]. In human NAFLD, activation of AMPK inhibited acetyl-CoA carboxylase (ACC), which in turn appeared to reduce hepatic lipid storage [377]. However, this finding is controversial, as AMPK does not appear to impact on lipid storage under obesogenic conditions [378]. However, this raises the possibility that, the findings in this chapter, related to AICAR-driven reduction in macrovesicular steatosis arise from AMPK activation. The partial restoration of macrovesicular steatosis observed following incubation with exogenous fumarate suggests that this is not the case, but it is reasonable to speculate that reductions in macrovesicular

steatosis in response to AICAR result from combined AMPK activation and PNC inhibition.

In addition to its impact on lipid accumulation, AMPK down-regulates the expression of NAD(P)H oxidase genes [379], which may alleviate oxidative stress in lipid-laden hepatocytes. This, in turn, could interfere with TCA cycle and ETC activity. However, as shown in Chapter 3, many of these genes were already downregulated, suggesting that AICAR-mediated AMPK activation does not impact on the isotopic tracing studies presented in this chapter. However, this cannot be fully ruled out without further investigation.

4.6 Conclusion

Taken together, I demonstrate for the first time that the development of steatosis in HLCs associates with truncation of the TCA cycle. I further demonstrate that following exposure to LPO, mitochondrial metabolic pathways are re-wired in order to maintain TCA cycle activity, through PNC-driven fumarate accumulation, which is associated with the development of macrovesicular steatosis. These findings reveal a novel mechanism linked to hepatic steatosis and may lead to further understanding of transcriptional and metabolic rewiring associated with NAFLD.

Chapter 5: Exploring links between altered TCA cycle activity and TET enzyme activity in steatotic cells

5.1 Introduction

The TCA cycle metabolites αKG, succinate and fumarate are allosteric regulators of the αKG-dependent dioxygenase family of enzymes; their activity is promoted by αKG and inhibited by succinate and fumarate [134–136]. The binding kinetics of substrates activating (Michaelis constant, Km)) or inhibiting (half maximal inhibitory concentration, IC_{50}) TET1 and TET2 enzyme activity are outlined in Tables 5.1 and 5.2, respectively. These values are currently unknown for TET3. This enzyme family include the Jumonji domain-containing histone demethylases (JHDMs), prolyl-4-hydroxylases, and the ten-eleven translocation (TET) dioxygenases [132,133]. The TET enzymes can regulate the epigenome through modulation of DNA methylation, by oxidising 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) [137,138]. Dysregulation of the TET enzymes may also be important in the pathogenesis of liver disease. It was previously observed that in murine and human models of NAFLD, there are alterations in 5hmC at specific loci, suggesting a role for the TET enzymes [128,208]. Studies in mouse models of HCC also suggest that the TET enzymes may contribute to the progression of liver cancer, through perturbation of normal DNA demethylation events at promoters [247].

In this chapter, I aimed to understand whether there is an association between perturbed TCA cycle activity in steatotic HLCs and TET enzyme activity. I performed these studies in the hPSC-derived NAFLD model used throughout this thesis [128].

Compound	TET1	TET2
Fe^{2+} (μM)	4.8 ± 4	125 ± 85
αKG (μM)	55 ± 20	3.6 ± 3
Oxygen (μM)	30 ± 10	30 ± 3

Table 5.1. Michaelis constant (Km) values for binding of substrates promoting TET1 and TET2 enzymatic activity. Adapted from Laukka *et al* (2016) [133].

Compound	TET1	TET2
Succinate (μM)	540 ± 100	570 ± 90
Fumarate (μM)	390 ± 160	400 ± 70

Table 5.2. The IC_{50} values for inhibition of TET1 and TET2 enzymatic activity. Adapted from Laukka *et al* (2016) [133].

5.1.1 Hypothesis

Steatosis-associated fumarate accumulation inhibits TET enzyme activity, leading to alterations in global levels of 5hmC.

5.1.2 Aims

1. To characterise changes in global levels of 5mC and 5hmC in response to development of steatosis in HLCs.
2. To determine whether changes in 5hmC were associated with altered expression of the TET enzymes, or alteration of their function.
3. To explore associations between altered genic 5hmC and gene transcription.

5.2 Materials and methods

5.2.1 Differentiation of hPSCs to HLCs and induction of lipid accumulation

Human female H9 pluripotent stem cells (hPSCs) were differentiated to hepatocyte-like cells (HLCs), and lipid accumulation induced as described in Chapter 2, section 2.1-2.2.

5.2.2 Ultra-performance liquid chromatography (UPLC) analysis of global 5mC and 5hmC

Sample preparation

2.5µg DNA in 43µL water, per sample, was incubated at 95 °C for 10 minutes. Then, 5µL of 10x T7 DNA polymerase reaction buffer and 10 U T7 DNA polymerase (Thermo Scientific, Renfrew, UK) was added. The reaction was incubated overnight at 37 °C, followed by incubation at 75 °C for 10 minutes. Finally, samples were centrifuged for 45 minutes at > 12,000 x g, before aspirating the supernatant and drying down using a Speed Vac. Samples were stored at -80 °C until time of analysis.

UPLC

Samples were analysed by Dr Jimi Wills at the mass spectrometry facility, Institute for Genetic and Molecular Medicine, University of Edinburgh, as described in Chapter 2, section 2.6.3.

5.2.3 DNA immunoprecipitation (DIP-seq) analysis of 5hmC enrichment

DNA was fragmented using a Covaris® E220 Focused-Ultrasonicator (Covaris, Brighton, UK), before immunoprecipitating hydroxymethylated fragments. Immunoprecipitated DNA was then sent to the Wellcome Trust Clinical Research Facility (Western General Hospital, Edinburgh), where it underwent sequencing. I analysed the sequencing data with the assistance of Sara Wernig-Zorc (University of

Regensburg, Germany), using the MACS2 and Homer pipelines. Full details of DNA preparation and bioinformatic analyses are located in Chapter 2, section 2.6 and 2.7.

5.2.4 3D-RNA seq analysis of alternative splicing

The RNA-seq dataset generated in Chapter 3 was re-analysed to explore alternative splicing events occurring in steatotic HLCs. Instead of using FeatureCounts, transcripts were quantified using Kallisto [380]. Transcript counts were then loaded into the 3D RNA-seq web-based app, with the assistance of Dr Wenbin Guo [381].

5.3 Results

5.3.1 Macrovesicular steatosis is associated with global changes in 5hmC, but not in expression of the TET enzymes

As fumarate is a known inhibitor of the TET enzymes, I wanted to explore whether increased abundance impacted on global levels of 5mC and 5hmC. UPLC analysis revealed a significant increase in global 5hmC (**Figure 5.1**) in steatotic HLCs. To determine whether this arose as a result of altered TET enzyme expression, I investigated expression of TET1, TET2, and TET3 in the RNA-seq analysis data generated in Chapter 3. This revealed that the presence of steatosis in HLCs is not associated with altered expression of these genes (**Figure 5.2**).

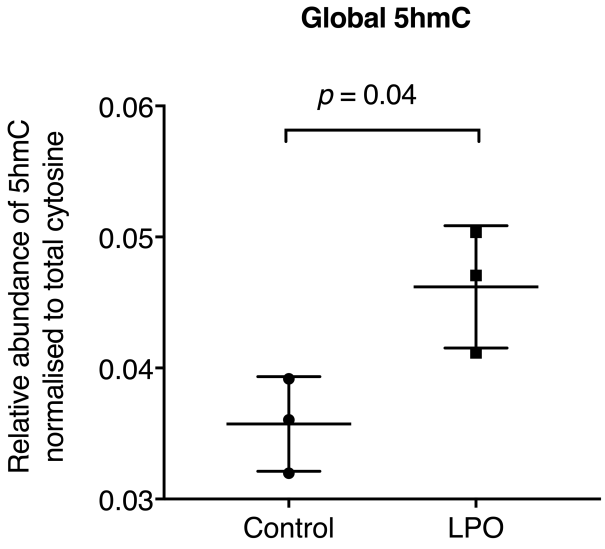


Figure 5.1 UPLC analysis reveals increased global 5hmC in steatotic HLCs. Global levels of 5hmC were increased in steatotic HLCs. For each group, $n = 3$. Data were analysed using a two-tailed Student's t-test. Data are expressed as mean ± SD.

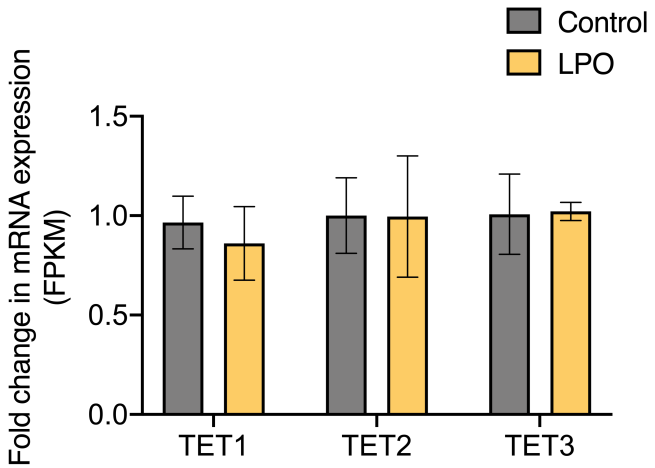


Figure 5.2. TET enzyme expression is unaltered in steatotic HLCs. Transcriptomic analysis performed in Chapter 3 shows that TET mRNA expression is unaltered in steatotic HLCs. For each group, *n* = 3. Data were analysed using a two-tailed Student's t-test. Data are expressed as mean ± SD.

Although expression of the TET enzymes was unchanged in steatotic HLCs, I questioned whether increased fumarate accumulation would have a functional impact on their activity, as determined by changes in 5hmC. To test this, genome-wide 5hmC sequencing was performed, identifying 3294 differentially hydroxymethylated regions (DHRs) between LPO exposed and control cells (>2-fold change). Within each group, samples were highly correlated, suggesting that the 5hmC patterns are stable between replicate samples (**Figure 5.3**), with the majority of changes located in intergenic (43%) or intronic (51%) regions (**Figure 5.4**). The remaining alterations in 5hmC enrichment were located in the 3' (0.7%) or 5' (0.1%) UTR sites, exons (1.3%), intergenic regions (43%), introns (51%), non-coding regions (0.4%), promoters (1.7%), or TSSs (1.1%).

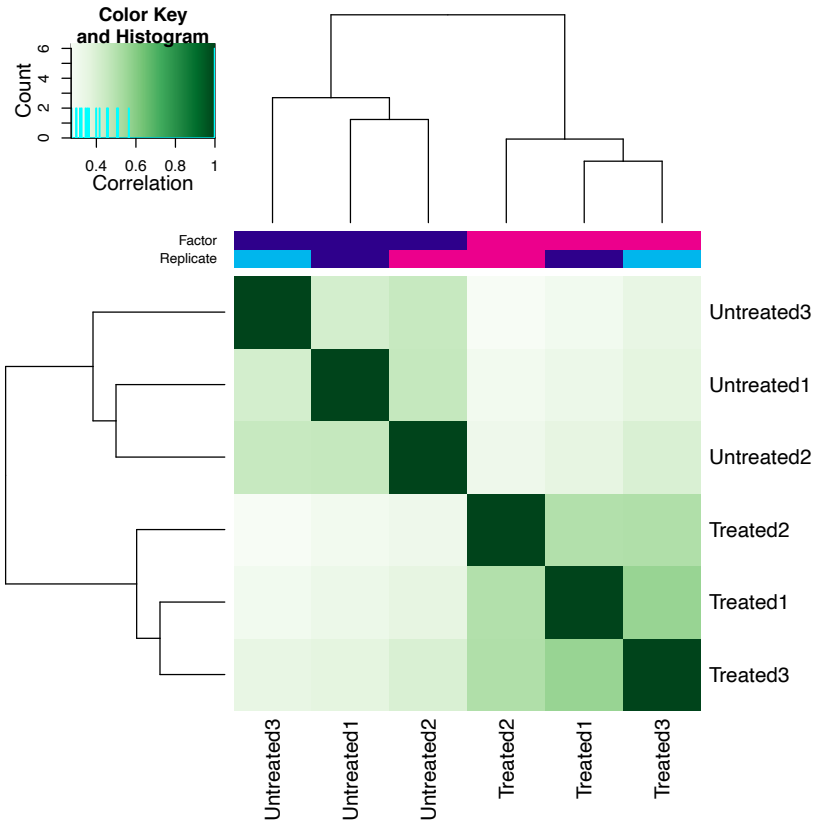


Figure 5.3 Correlation heatmap of control vs. LPO groups, following hmeDIP sequencing, clustered by Euclidean distance. Untreated = control group. Treated = LPO-treated group. For both control and LPO, n = 3 biological replicates/group. Figure is adapted from Sinton *et al*, 2020 [304]

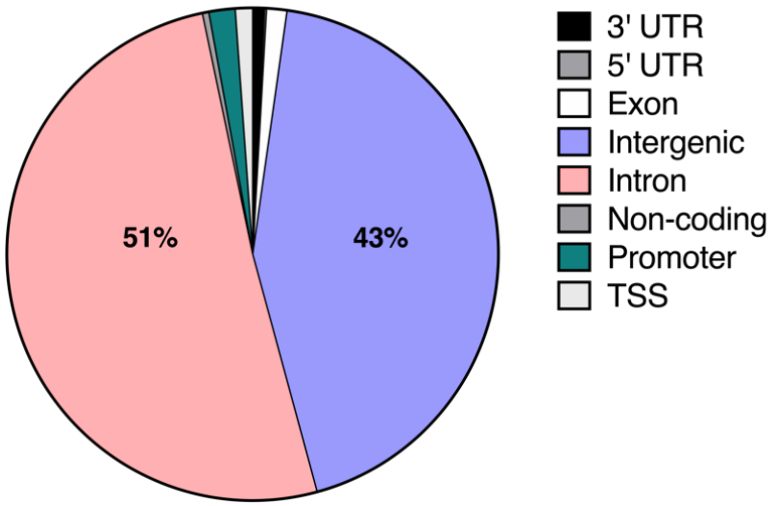


Figure 5.4. Proportion of DHRs associated with different regions of the genome, showing the majority located in intrergenic and intronic regions. Regions were differentially enriched/depleted in LPO-treated HLCs compared with controls. For both control and LPO, n = 3 biological replicates/group. Figure is adapted from Sinton *et al*, 2020 [304]

5.3.2 5hmC alterations in steatotic HLCs occurs in genic regions associated with tissue remodelling and Wnt/beta-catenin signalling

To further explore the changes in 5hmC in response to steatosis in HLCs, I performed a STRING analysis. Similar to pathway enrichment, this analysis can determine relationships between clusters of genes that are differentially expressed and offers a visual representation of this. It also enables identification of putative hub genes, which have high connectivity to other genes. These may represent key genes that are involved in the disruption of pathways associated with disease [382]. Performing this analysis on genic regions identified in section 5.3.1 generated a map consisting of 5 clusters of genes (**Figure 5.5**). Clusters related to Wnt/beta-catenin signalling (yellow), PI3K signalling (green), collagen deposition (aqua), or DNA damage repair (blue). Genes represented by red spheres did not cluster. Within the yellow cluster, I also identified CTNNB1, which encodes beta-catenin, as a hub gene. CTNNB1 is a

key component of the Wnt/beta-catenin signalling pathway and is highly connected to other genes within this cluster, suggesting that it may play a key role in NAFLD pathogenesis. DIP-seq analysis revealed a 3.7-fold decrease in 5hmC in intron 33 of the CTNNB1 gene, correlating with a -0.6-fold decrease in mRNA expression (Chapter 8, Appendix 5). Within the aqua cluster are several genes related to tissue remodelling, including collagen, type VI, alpha 1 (COL6A1), a gene known to be regulated by 5mC [383]. DIP-seq analysis of COL6A1 revealed a 3.1-fold increase in 5hmC in the gene promoter as well as a 4.1-fold increase in 5hmC in intron 3, concomitant with a 1.2-fold increase in mRNA expression (**Supplementary file MCS1 – DEGs in LPO-treated HLCs**). STRING analysis also revealed enrichment of the UniProt keyword term KW0025: alternative splicing, highlighting that a significant proportion of highlighted genes are subjected to this process and have multiple isoforms.

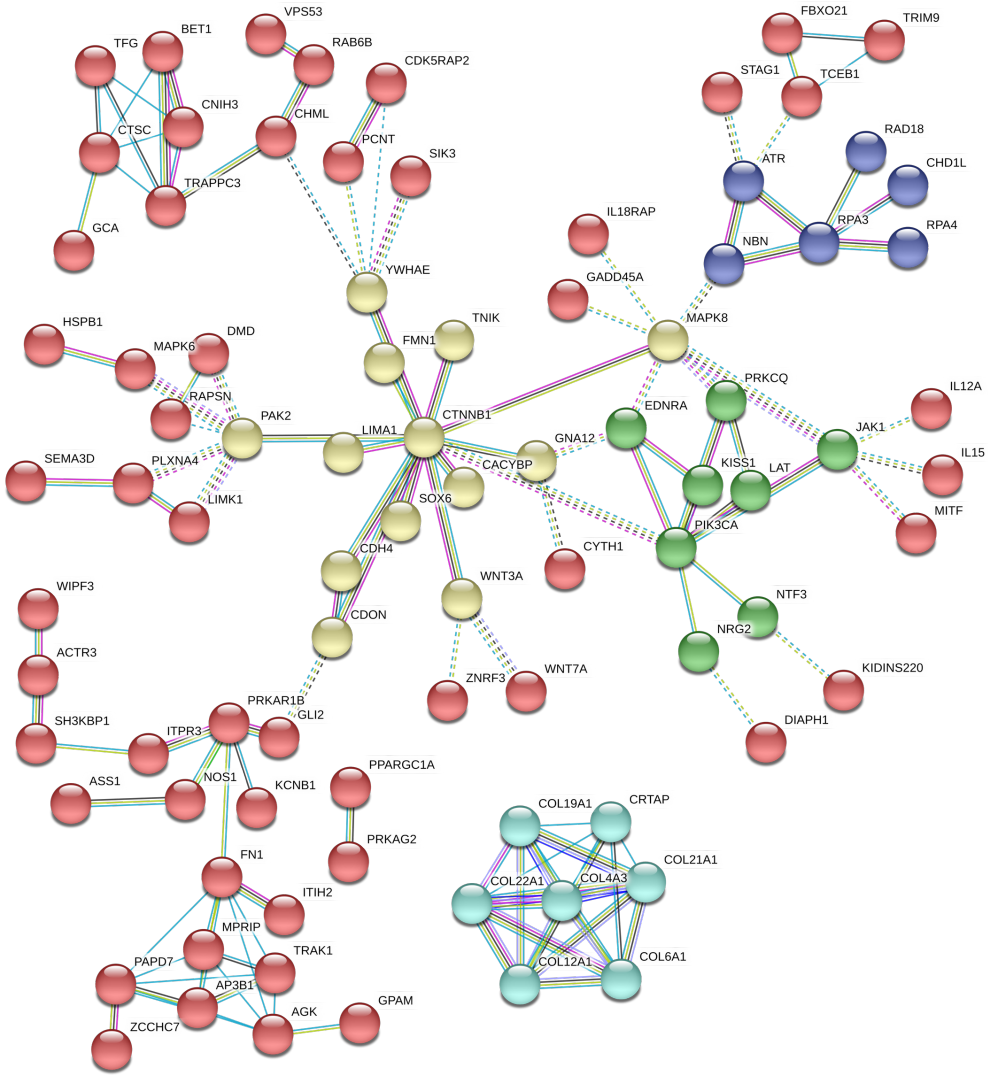


Figure 5.5. STRING analysis of DHRs in genic regions. Analysed genic regions were exons, introns, and transcriptional start sites, regardless of fold change in enrichment. Clusters were calculated using k-means clustering. Mapped genes had the highest confidence interaction score (0.9). Line colours represent the following: blue – information from curated databases; pink – interaction confirmed experimentally; yellow – information from text mining; black – genes known to be co-expressed; green – genes are in the same gene neighbourhood. Dashed lines occur where different clusters are linked. Each sphere colour represents a cluster.

5.3.3 Differential 5hmC enrichment in promoter regions shows a moderate negative correlation with mRNA expression in steatotic HLCs

It is thought that 5hmC in gene body regions may functionally relate to mRNA transcription [247]. Therefore, I generated heatmaps representing mean changes upstream of and across the TSS, with cluster 1 showing subtle increases in 5hmC within the upstream region (**Figure 5.6**). Sliding window analysis of 5hmC upstream of and across the gene body supports this finding, revealing subtle increases in 5hmC within putative promoter regions following LPO exposure (**Figure 5.7**).

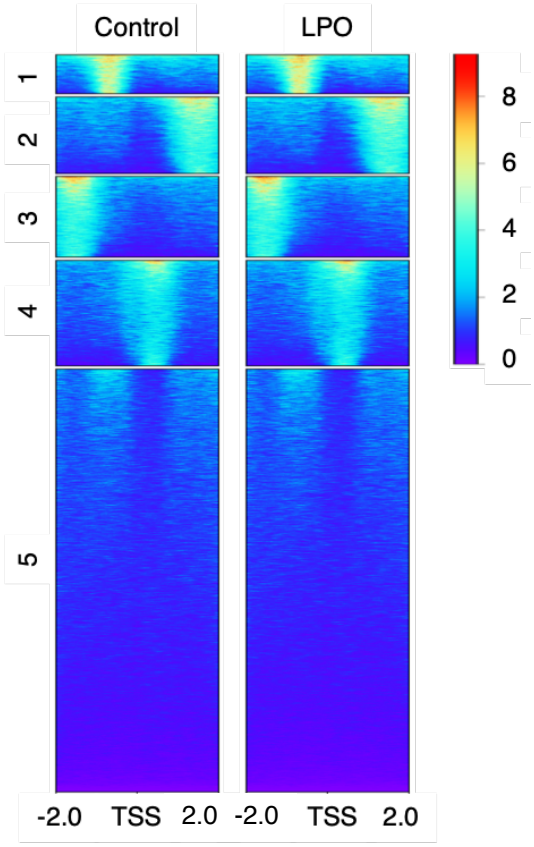


Figure 5.6. Heatmap analysis of DHRs surrounding transcriptional start sites highlights subtle increases in 5hmC in putative gene promoter regions. Heatmap was generated by k-means clustering, with an FDR of <0.05. Regions were differentially enriched/depleted in LPO-treated HLCs compared with controls. For both control and LPO, n = 3 biological replicates/group. Numbers on the y-axis represent clusters. TSS = transcriptional start site. Figure is adapted from Sinton *et al*, 2020 [304].

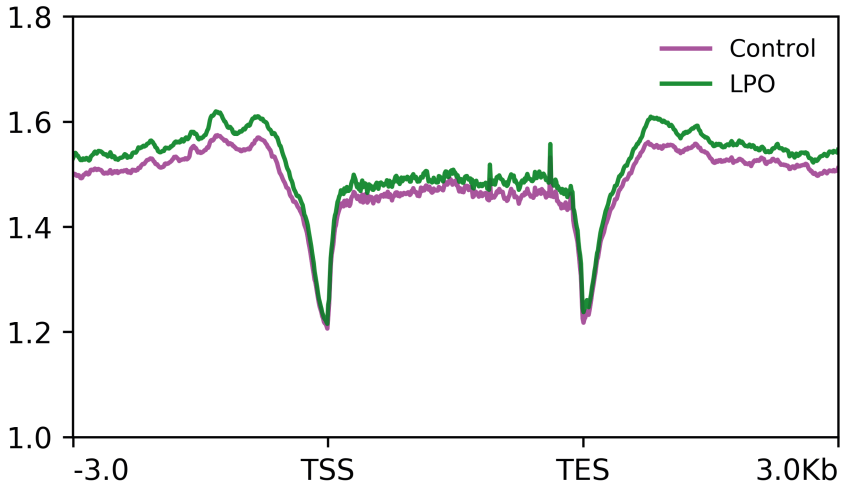


Figure 5.7. Sliding window analysis of average 5hmC levels of the transcriptional start site (TSS), gene body, and transcriptional end site (TES) shows minimal changes between control and LPO groups. Regions were differentially enriched/depleted in LPO-treated HLCs compared with controls. For both control and LPO, n = 3 biological replicates/group. Figure is adapted from Sinton *et al*, 2020 [304]

In line with this, I integrated DHR and RNA-seq data to look at promoter regions upstream of the transcriptional start site. In doing so, I identified 12 promoter regions with differential enrichment of 5hmC (>2-fold) and where mRNA expression was altered >0.5-fold (**Table 5.1**). Linear regression of these regions identified a moderate but significant negative relationship between 5hmC enrichment and mRNA expression (**Figure 5.8**). Of these genes 7 (EPHX3, ERO1B, CSGALNACT1, DOC2A, COL6A1, CASP1 and TMEM88) were previously shown to be dysregulated in the pathogenesis of NAFLD or progression to cirrhosis [384–388]. This may indicate a role for 5hmC in the regulation of these genes in NAFLD pathogenesis.

Ensembl ID	Gene symbol	mRNA expression fold change	5hmC fold change
ENSG00000181418	DDN	3.57	-2.84
ENSG00000105131	EPHX3	2.38	-3.82
ENSG00000086619	ERO1B	1.65	4.05
ENSG00000147408	CSGALNACT1	1.33	3.09
ENSG00000149927	DOC2A	1.26	-2.49
ENSG00000142156	COL6A1	1.24	3.13
ENSG00000137752	CASP1	0.98	2.44
ENSG00000119900	OGFRL1	0.71	2.21
ENSG00000145247	OCIAD2	0.71	3.02
ENSG00000183044	ABAT	-1.26	3.04
ENSG00000144395	CCDC150	-1.50	3.36
ENSG00000167874	TMEM88	-2.02	3.60

Table 5.1. Promoter regions of genes with both altered mRNA expression and enrichment of 5hmC. Table is taken from Sinton *et al*, 2020 [304].

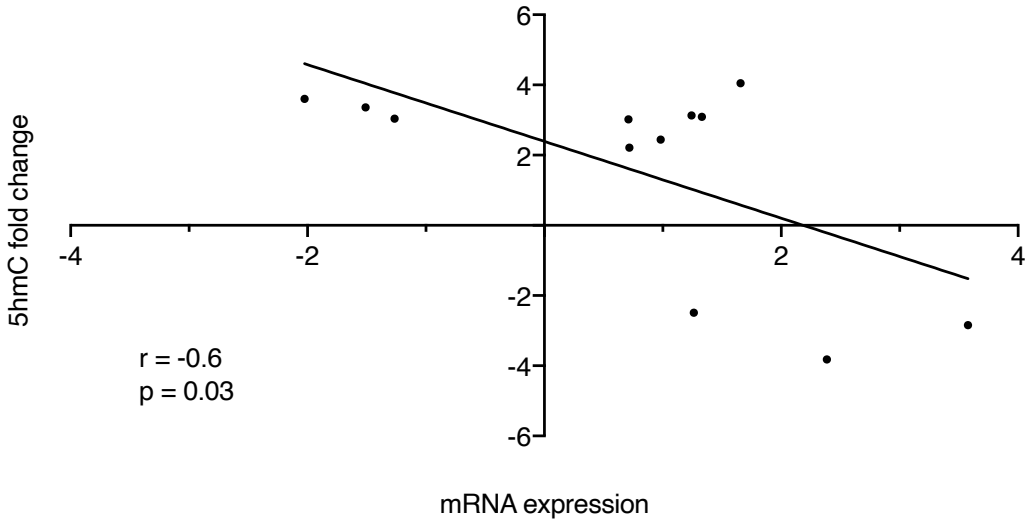


Figure 5.8. Steatotic HLCs show a moderate negative correlation between putative promoter 5hmC enrichment and mRNA expression. The cut-off for inclusion was: 5hmC enrichment/depletion >2.0; mRNA expression log_2 fold change >0.5. Data were analysed by linear regression. Figure is adapted from Sinton *et al*, 2020 [304]

5.3.4 In steatotic HLCs, decreased 5hmC in the alternative promoter of PPARGC1A is associated with altered isoform expression

As 5hmC deposition is thought to influence alternative splicing, [205,389], and this was an enriched pathway in the STRING analysis, I was interested in understanding whether 5hmC enrichment in genic regions in steatotic HLCs was associated with differential splicing events. To this end, the RNA-seq dataset generated in Chapter 3 was reanalysed using the 3DRNA-seq analysis pipeline [381], and then integrated with the DIP-seq data presented in this chapter. This analysis identified that depletion of 5hmC in the alternative promoter of PPARγ coactivator-1-α (PPARGC1A; **Figure 5.9**) is associated with increased expression and differential transcript usage of transcript variant 2 of this gene.

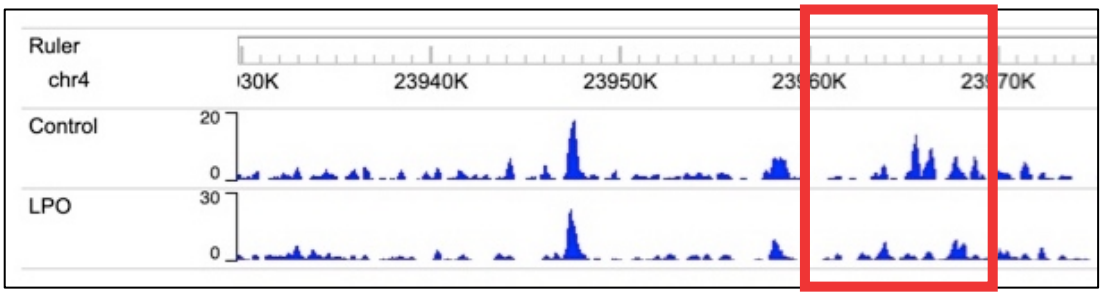


Figure 5.9. Steatosis in HLCs is associated with depletion of 5hmC in the alternative promoter of PPARGC1A.

	5hmC fold change	mRNA fold change	Differential transcript usage
Fold change	-3.73	+1.52	+1.89
Transcript	Intron 2. Nearest promoter: NM_013621, transcript variant 2	No data	NM_0132621.5, corresponds to transcript variant 2

Table 5.2 Depleted 5hmC in the alternative promoter of PPARGC1A is associated with increased expression of transcript variant 2.

PPARGC1A can be spliced to multiple isoforms, and although the function of these different isoforms remains unclear, when translated they are each subject to different PTMs, indicating that each has a distinct function (**Figure 5.10**) [390].

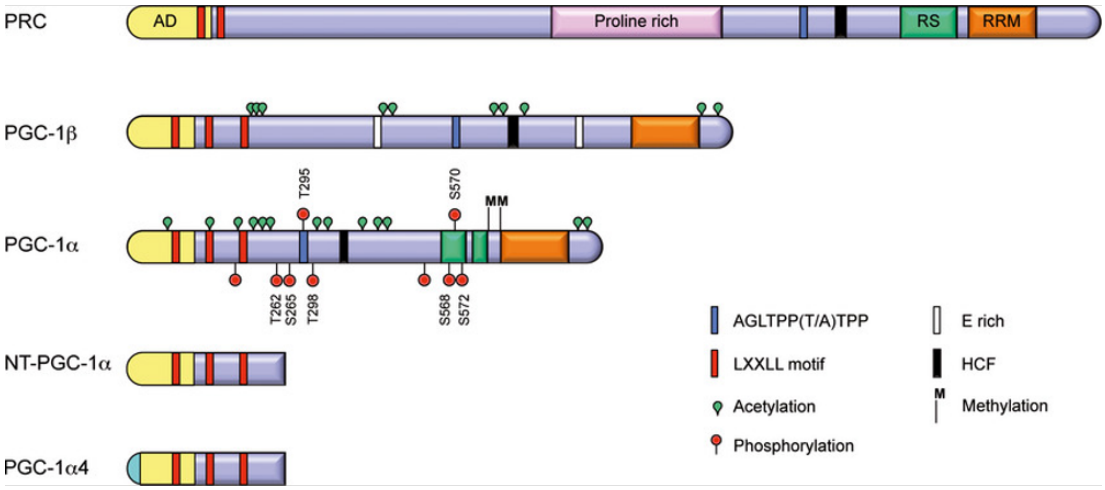


Figure 5.10. Domain structures of PGC1-A (protein name for PPARGC1A). This schematic shows the full length PGC1-A (PRC), as well as the shorter isoforms PGC-1β (encoded by PPARGC1A 1b), PGC-1α (encoded by PPARGC1A 1a), NT-PGC-1α (encoded by NT-PPARGC1A) and PGC-1α4 (encoded by PPARGC1A 1c). Phosphorylation, acetylation and methylation are the most common PTMs identified for PGC-1α and PGC-1β. AD = activation domain; RS = serine/arginine-rich stretch; HCF = host factor cell; LXLL = leucine-rich motif. Figure taken from Villena (2014) [390].

To further interrogate this, I performed RT-qPCR of total PPARGC1A and then of multiple isoforms generated from alternative splicing (**Figure 5.10**). This analysis confirmed that total PPARGC1A expression increases in steatotic HLCs, and that there is an increase in the intron 2-derived PPARGC1A-c, as well as the truncated NT-PPARGC1A isoform. Taken together, this suggests that 5hmC depletion in intron 2 of PPARGC1A may impact alternative splicing of this gene, leading to preferential transcript usage of PPARGC1-c. However, the significance of this is unclear.

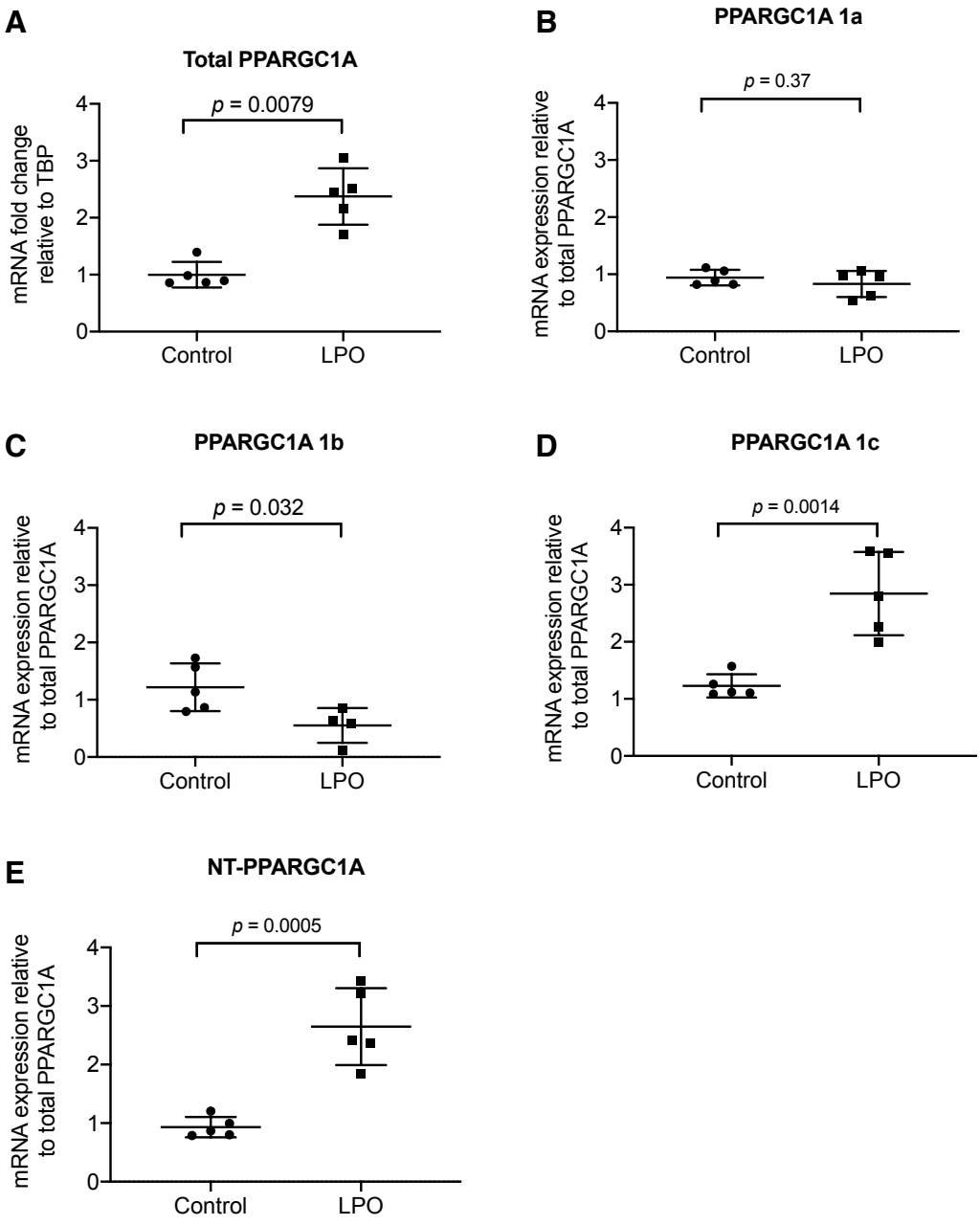


Figure 5.11. Steatosis in HLCs is associated with differential splicing of PPARCG1A. (**A**) Total PPARGC1A increased in steatotic HLCs. (**B**) Exon 1a derived transcript (PPARGC1A 1a) was unchanged in steatotic HLCs. (**C**) Exon 1b-derived transcripts (PPARGC1A 1b) decreased. In contrast, (**D**) exon 1c-derived transcripts (PPARGC1A 1c) and (**E**) truncated transcript (NT-PPARGC1A) expression increased. Data analysed using two-tailed Student t-test and expressed as mean ± SD, n=5 biological replicates/group.

5.4 Discussion

5.4.1 Macrovesicular steatosis is linked to altered TET enzyme activity, leading to global changes in enrichment of 5hmC

A potential consequence of fumarate accumulation in steatotic HLCs is inhibition of αKG-dependent dioxygenase enzyme activity. Given the previous data showing altered 5hmC in a mouse model of NAFLD and the potential importance of the TET enzymes in HCC [208,247], I used a DIP-seq approach to measure changes in 5hmC across the genome. 5hmC is a stable epigenetic modification that may influence transcription, and while the nature of this influence is unclear, studies suggest that 5hmC enrichment at transcriptional start sites (TSSs) leads to repression of transcription [204]. Recent studies in mice showed that 5hmC enrichment is reversibly altered at specific loci in response to fat accumulation in the liver, indicating that the TET enzymes may play a role in NAFLD pathogenesis [208]. I observed widescale changes in 5hmC enrichment, predominantly in intronic and intergenic regions of the genome. There are studies suggesting that enrichment of intronic 5hmC may influence transcription in hepatocyte differentiation [391]. Furthermore, studies of mouse dilated cardiomyopathy demonstrated an association between the disease and enrichment of intronic 5hmC [392]. One possibility is that the influence of 5hmC within these regions is related to long range chromatin interactions, which was shown to have a greater impact on transcription than enhancer enrichment in differentiation of human $CD4^+$ T cells [393]. Integrating the data presented here with a chromatin conformation capture approach, to explore interaction of 5hmC with distal regulatory elements may provide a clearer insight into whether intergenic or intronic 5hmC influences transcription in steatotic HLCs.

5.4.2 Altered enrichment of 5hmC is found in genes related to Wnt/beta-catenin signalling and tissue remodelling in steatotic HLCs

When examining 5hmC across gene bodies, STRING analysis revealed alterations in enrichment in genes related to tissue remodelling, and Wnt/beta-catenin signalling. As Wnt/beta-catenin signalling is known to induce expression of genes related to

tissue repair and remodelling [394], it is notable that both of these pathways were enriched in this analysis. Wnt/beta-catenin signalling is increased in ~50% of patients who present with NAFLD-associated HCC [395]. Furthermore, a significant proportion of these patients do not have mutations in CTNNB1, the gene encoding beta-catenin, suggesting that expression may be modulated by the epigenome [396]. Previous studies in mouse models of HCC have shown a gain of 5mC with a concomitant loss of 5hmC in the promoter core of CTNNB1 [247]. Further to this, histone deacetylase 8 (HDAC8) regulates CTNNB1 in NAFLD-associated HCC [397]. However, it is unclear which loci in CTNNB1 undergo HDAC8-mediated deacetylation. The presence of either 5mC or 5hmC can recruit MeCP2 [398], which, in turn, can recruit HDACs [399] or histone methyltransferases [400] to the DNA. In steatotic HLCs it is unclear whether loss of 5hmC in CTNNB1 influences transcription of the gene, particularly given the location of the depletion. Furthermore, it is unclear whether depletion of 5hmC here is a result of increased TET enzyme activity, further oxidising it to downstream derivatives, or whether there is decreased activity, leading to increased levels of 5mC.

In addition to these changes, a number of genes related to tissue remodelling displayed altered enrichment of 5hmC, with changes mostly occurring in intronic regions. These included collagen, type VI, alpha 1 (COL6A1) and collagen, type XII, alpha 1 (COL12A1). Upregulation of these genes in NAFLD is typically associated with stellate cell activation in response to inflammation, leading to development of hepatic fibrosis [387]. In steatotic HLCs, there was an upregulation of COL6A1 and COL12A1, which are known to be involved in development of NAFLD-related cirrhosis in humans [387,401]. Previous studies have shown that 5mC in the promoter of COL6A1 can influence transcription of the gene [383]. However, this was determined using bisulphite sequencing, which is unable to discriminate between 5mC and 5hmC [402]. Therefore, it is unclear which of these DNA modifications, in this context, influences transcription. Whereas it appears that COL6A1 expression may be influenced by TET activity, this is less clear in the case of COL12A1, where increased levels of 5hmC occur in a late intron within the gene body. As both Wnt/beta-catenin signalling and collagen deposition are associated in NAFLD, the data presented here

suggest that altered enrichment of genic 5hmC may play a role in disruption of these pathways. However, further studies are required to confirm or refute this.

5.4.3 There is a moderate negative correlation between 5hmC and mRNA expression associated with steatosis in HLCs

Changes in 5hmC at specific loci did not, in general, correlate with transcriptional changes in steatotic cells. However, as gene promoter 5hmC has previously been correlated with transcriptional perturbation [247], I went on to explore this relationship in steatotic HLCs. This analysis revealed a moderate negative correlation between 5hmC enrichment and mRNA expression in 12 genes, 7 of which were previously shown to be associated with NAFLD: EPHX3, ERO1B, CSGALNACT1, DOC2A, COL6A1, CASP1 and TMEM88 [384–388].

ERO1B is an endoplasmic reticulum (ER) oxidoreductase, which catalyses the formation of sulphide bonds [403]. Under ER stress, the unfolded protein response (UPR) is activated, which can induce increased expression of ERO1B [404]. There is evidence to suggest that LPO may induce ER stress [42], a known contributor to NAFLD [44]. Further to this, the UPR can induce activation of the NLRP3 (NOD-, LRR- and pyrin domain-containing protein 3) inflammasome [405], and downstream activation of CASP1 (caspase-1), leading to apoptosis [406]. Upregulation of both ERO1B and CASP1 suggests that steatotic HLCs are undergoing increasing levels of apoptosis, which is thought to be involved in the progression of benign steatosis to NASH [407].

Upregulation of EPHX3 (epoxide hydrolase 3) mRNA expression correlated with a loss of 5hmC in the promoter of this gene. Whilst the precise role of EPHX3 in NAFLD has not been described, activation of this enzyme is thought to be beneficial for the treatment of NASH, and has been shown to improve lipid homeostasis in hepatocytes [408]. However, this study was performed in mice using a methionine-choline-

deficient diet model of NASH, which is poorly representative of the disease in humans [238]. Upregulation of EPHX3 in steatotic HLCs may be a protective mechanism, assisting in maintenance of lipid homeostasis. However, knockout and overexpression studies would be required to further interrogate this.

In steatotic HLCs, TMEM88 (transmembrane protein 88) mRNA expression was downregulated, with a concomitant increase in promoter 5hmC. TMEM88 is known to inhibit Wnt/beta-catenin signalling [409] and, therefore, its downregulation in steatotic HLCs may de-repress this signalling pathway. As discussed in section 5.4.2, Wnt/beta-catenin signalling is increased in ~50% of patients who present with NAFLD-associated HCC [395]. Previous studies have characterised methylation of the TMEM88 promoter, associating it with unfavourable prognosis in non-small cell lung cancer [410]. This study used bisulphite sequencing to determine the presence of methylation in the TMEM88 promoter, which, as discussed in section 5.4.2 is unable to discriminate between 5mC and 5hmC [402]. Therefore, the authors of this study may have been observing enrichment of promoter 5hmC rather than 5mC, as observed in this study. Taken together, repression of TMEM88 expression by promoter 5hmC enrichment may de-repress Wnt/beta-catenin signalling in steatotic HLCs and may be an indicator of downstream development of NAFLD-related HCC.

5.4.4 Depletion of 5hmC in the PPARGC1A alternative promoter correlates with increased expression of specific isoforms

Steatosis in HLCs was associated with depletion of 5hmC in the promoter of PPARGC1A, with corresponding increases in mRNA expression of exon 1c-derived transcripts (PPARGC1A 1c) and truncated PPARGC1A (NT-PPARGC1A). The canonical PPARGC1A promoter is known to be subject to epigenetic modification, with studies showing increased enrichment of 5hmC in response to exercise in mice [411]. However, the same study observed no changes in 5hmC in the alternative promoter, suggesting that hydroxymethylation of this promoter is context dependent. Whilst changes in canonical promoter methylation corresponds with changes in gene

transcription, it is unknown whether this is the case for the alternative promoter, and further studies are required to determine this. A similar study showed increased expression of alternative promoter-derived PPARGC1A isoforms and found that PPARGC1A 1c had a similar function to canonical promoter-derived isoforms [412]. Therefore, further studies are required to understand whether preferential expression of alternative promoter-derived isoforms of PPARGC1A is of functional importance in NAFLD.

5.5 Experimental limitations

Although this study focused on promoter 5hmC enrichment, it possible that differential 5hmC enrichment is occurring in regions associated with expression of enhancers, non-coding RNAs, or microRNAs, which in turn may influence mRNA transcription [413–415]. Further to this, the DIP-seq approach utilised here does not provide base-pair resolution. Using combined bisulphite and oxidative bisulphite sequencing would provide this level of resolution, and may give additional insight into the interplay between 5hmC enrichment and transcriptional regulation [416].

I hypothesised that steatosis-associated fumarate accumulation would influence TET enzyme activity, as demonstrated by altered 5hmC enrichment. However, as previously mentioned, fumarate is an allosteric inhibitor of the whole dioxygenase family of enzymes, which may impact on transcriptional regulation in NAFLD, including effects on HIF1α or the histone demethylase enzymes [417,418]. Further to this, in steatotic HLCs there was also an increase in αKG, which is an activator of the TET enzymes [137]. It is suggested that as αKG has a higher binding affinity for the dioxygenases that it will always outcompete fumarate for binding [209]. From the data presented here, it is unclear whether changes in 5hmC arise from inhibition of TET activity, leading to decreased conversion of 5mC to 5hmC, or increased activity, leading to increased oxidation of 5hmC to 5fC.

It is important to note that the data presented in this chapter are correlative, and do not show a direct link between 5hmC and mRNA expression. In this regard, the negative correlation between promoter 5hmC and mRNA expression was based on a very limited number of genes raising the question of whether this is functionally relevant. Further mechanistic studies, including knockout and overexpression experiments, may provide further insight into the functional relevance of correlations between 5hmC and mRNA expression in steatosis.

5.6 Conclusion

This genome-wide study of 5hmC enrichment in steatotic HLCs provides a novel insight into the epigenetic regulation of NAFLD-related genes. Here, I reveal genome-wide alterations in 5hmC, and a potential correlation between promoter 5hmC enrichment and transcription of genes relevant to NAFLD. Furthermore, this study further supports that Wnt/beta-catenin signalling is important in NAFLD and suggests that it is influenced by overnutrition-mediated epigenetic regulation. This study opens the door to further exploration of how environmental cues influence the epigenome and transcriptional regulation in NAFLD.

Chapter 6: Discussion

6.1 Summary

NAFLD is the most prevalent form of hepatic disease globally, representing an extraordinary burden on healthcare systems. There is currently a lack of non-invasive diagnostics or therapeutics available to reverse pathology, which is a reflection of the complexity of this disease [15]. As global obesity rates soar, with concurrent increases in the prevalence of NAFLD, the need for easy diagnostics and therapeutic interventions is becoming ever more urgent. The work in this thesis explores, at high resolution, the impact that overnutrition has on hepatic mitochondrial metabolism in a human-relevant hPSC model of NAFLD. In this thesis I show, for the first time, that rewiring of the TCA cycle directly impacts on lipid accumulation in hepatocytes, presenting a possible area of interest for the development of novel therapeutics. Further to this, I also demonstrate that these changes are associated with transcriptomic and epigenomic changes, which may be important in the development of NAFLD.

6.2 Novel findings

6.2.1 Hepatic steatosis is associated with impaired mitochondrial respiration

Previous studies have demonstrated alterations in mitochondrial respiration associated with NAFLD [34,130]. In 2015, Koliaki *et al* showed that hepatic mitochondria from NAFLD patients displayed increased maximal respiration in the absence of changes in mitochondrial content [130]. In contrast, they identified increased hepatic mitochondrial content in NASH patients, which was associated with decreased maximal respiration. Based on these findings they hypothesised that mitochondria in steatotic hepatocytes can adapt to cellular stress but that this is lost when the disease progresses to NASH. However, there were a number of limitations to this study, potentially confounding their findings. The authors studied whole tissue and the study, therefore, does not account for cell type-specific changes in mitochondrial respiration. Different cell types have varying metabolic profiles depending on their energetic needs. For example, activated Kupffer cells, which are

found in the liver during NASH, are more glycolytic [419], whereas steatotic hepatocytes appear to rely more on oxidative metabolism [34]. Whilst this could be overcome by using a cell sorting approach, the metabolome is very sensitive to changing environmental stimuli, and this process has been shown to alter cellular redox status [253]. The use of a single cell type, as described in this thesis, overcomes these challenges, enables the study of the metabolome without introducing environmental stressors or the confounding effects of multiple cell populations with varying metabolic profiles.

In addition to this, the authors of this study used isolated mitochondria to examine the respiratory response to NAFLD. Mitochondrial isolation is understood to impact on organelle function as it can result in fragmentation, impairing respiratory function [420]. By culturing and differentiating hPSCs *in situ*, I was able to measure mitochondrial respiration without this confounding effect. Another limitation of the Koliaki study was that the authors only analysed expression of genes associated with mitochondrial biogenesis and did not perform transcriptomic analyses to determine whether there were NAFLD-associated changes in transcripts encoding respiratory complexes. The transcriptomic analyses performed in the studies presented here demonstrate that hepatic steatosis is associated with decreased expression of genes related to the TCA cycle and ETC. Whilst this does not demonstrate cause, it suggests that downregulation of these genes plays a role in the impairment of mitochondrial respiration.

In this thesis, I demonstrate that steatotic HLCs have decreased maximal respiration capacity, but in the absence of changing mitochondrial content. This suggests that this model may reflect a transition phase, falling between the NAFLD and NASH phenotypes, and that impaired respiratory function may precede increased mitochondrial biogenesis.

6.2.2 Rewiring of the TCA cycle and subsequent fumarate accumulation drives macrovesicular steatosis

It was previously observed by Burgess *et al* that NAFLD in both rodents and humans appears to be associated with mitochondrial dysfunction and increased TCA cycle activity [34,109,129,421]. Whilst informative, studies in humans have been indirect, using blood plasma metabolites as a readout of TCA cycle activity [34]. However, the finding of increased TCA cycle activity in rodents, with increased levels of all TCA-cycle associated metabolites, support the findings in humans. This strongly indicates that hepatic TCA cycle function is disrupted in association with NAFLD.

Using the hPSC-based model of NAFLD described in this thesis, I was able to determine whether hepatic steatosis is indeed associated with TCA cycle disruption. Using stable isotope tracing studies, I found that steatosis is associated with TCA cycle truncation, inhibiting the conversion of succinate to fumarate. This correlates with the findings of Chapter 3, where I observed dysfunction in the ETC. This dysfunction most likely arises from inhibition of respiratory complex I, II, or III, but I was unable to determine which of these was responsible for decreased maximal respiration in mitochondria. Inhibited conversion of succinate to fumarate suggests that this is, in part, due to dysfunction of complex II, also known as succinate dehydrogenase. Further respirometry studies may provide further insight into this.

Steatotic HLCs were able to compensate for TCA cycle truncation by increasing purine nucleotide cycle (PNC) activity, leading to increased generation of fumarate. A similar compensatory mechanism was previously observed in cardiac ischaemia reperfusion injury [358], demonstrating that the TCA cycle is plastic and can adapt to changing metabolic demands. Further supporting this are studies of fumarate hydratase (FH), showing that knocking out the enzyme causes respiratory dysfunction but is not lethal, due to compensatory rewiring [127]. Notably, FH knockout resulted in fumarate accumulation, which, in turn, inhibited complex II activity. This raises the question of whether, in hepatic steatosis, increased PNC activity is compensating for TCA cycle truncation, or whether the truncation arises as a result of increased PNC

activity. Furthermore, this opens the question of whether the TCA cycle exhibits plasticity under homeostatic conditions or whether this is more generally associated with disease pathogenesis.

6.2.3 Macrovesicular steatosis is associated with moderate changes in enrichment of 5hmC in gene promoters

The association between transcriptional activity and 5hmC levels have previously been documented [206,212,222]. Furthermore, two studies reveal that changes in hepatic 5hmC are dynamic and can be induced by toxicological insults, with subsequent reversal upon withdrawal of the insult [422][423]. Studies of NAFLD have also identified dynamic changes in hepatic 5hmC, suggesting that it may play a role in the transcriptional changes observed in association with this disease [208]. As 5hmC levels are modulated by the TET enzymes, and the activity of these enzymes is inhibited by fumarate, I questioned whether increased fumarate would inhibit TET activity. Sequencing analyses revealed widespread changes in 5hmC enrichment in steatotic HLCs, with the majority located in intergenic or intronic regions. Whilst these regions could act as regulatory elements and influence the expression of distal genes, the data presented in this thesis do not provide evidence of this, and further studies are required. I also observed a moderate negative correlation between promoter 5hmC and gene transcription in 12 genes, as well as in the alternative promoter of PPARGC1A.

The data presented in this thesis support the idea that generation of 5hmC is highly dynamic and that the TET enzymes can respond rapidly to changing environmental cues, an essential property for transcriptional regulators. However, whether 5hmC is functionally important in the regulation of transcription and altered transcriptional programme associated with NAFLD remains unclear. The findings that 5hmC profiles are highly conserved, precisely regulated and tissue specific suggests a functional role [67,196,206,215]. Genetic manipulation of the TET enzymes in the model of NAFLD used in this thesis may provide further insight into their influence on the pathogenesis of this disease.

6.3 Limitations

This model of NAFLD utilises a single cell type cultured and differentiated in a monolayer. Whilst the model recapitulates the features of NAFLD that I aimed to study in this thesis, the model has limitations. As with many *in vitro* culture systems, these include the lack of other cell types and endocrine signals that cells receive *in vivo*. As described in Chapter 1, a major feature of NAFLD is the activation of Kupffer and stellate cells, which contribute to steatosis-associated inflammation and fibrosis. Furthermore, NAFLD in humans is generally associated with IR [424], which is not a feature of this model. However, it is also important to note that these limitations apply to other *in vitro* models of NAFLD and it is, therefore, important to interpret the work in the context of the considerable biological and technical limitations implicit in the field.

A further limitation of the studies is related to the study of 5hmC in steatotic HLCs. The use of DNA immunoprecipitation sequencing (DIP-seq) provides a highly reproducible method of profiling, and correlates well with qPCR analysis. However, despite the accuracy of sequencing in identifying regions of enrichment, it is unclear how well this methodology can detect smaller changes at occurring at typically highly enriched loci. Furthermore, this type of sequencing does not offer base-pair resolution. There are alternative techniques that offer this level of resolution, including oxidative bisulphite or TET-assisted bisulfite sequencing. These techniques have enormous potential but remain prohibitively expensive. Another advantage of bisulphite-based sequencing is the uniformity of bioinformatic pipelines used to analyses them. This is a further limitation of DIP-seq, where pipelines are variable, and generally based around the use of tools built for ChIP-seq. The lack of standardised pipelines makes it challenging to compare studies directly. Further to this, the studies presented here do not clarify whether NAFLD-associated changes in 5hmC enrichment arise from fumarate-induced inhibition of TET activity. It is possible that the changes observed arise because TET activity is inhibited and that 5mC cannot be oxidised to 5hmC and downstream derivatives. Alternatively, increased levels of αKG may promote increased TET activity, leading to increased conversion of 5mC to 5hmC and downstream derivatives. Future studies involving genetic

manipulation of individual TET enzymes, or their inhibition, will allow us to draw conclusions related to how altered mitochondrial metabolism influences their activity.

6.4 Future directions

The work in this thesis opens the door to a number of avenues of exploration. Delving further into the findings presented in this thesis, the next steps would include determining the mechanism by which fumarate accumulation induces macrovesicular steatosis. A number of approaches could help to address these questions, including further isotopic tracing analyses of non-polar metabolites to determine how TCA cycle metabolites and LPO are utilised to synthesise lipids. A further approach is to use mass spectrometry imaging, which can reveal the lipid species composition of lipid droplets at single-cell resolution [375].

A further question raised in this thesis is whether fumarate accumulation impacts TET enzyme activity. This is challenging as the TET enzymes are involved in differentiation [425], and as hPSCs differentiate they reach confluence, making them difficult to transfect. One strategy for knocking out TET function is to create a Cas9 knock-in, under a constitutive promoter, alongside guide RNAs under an albumin promoter. Overexpression is made more challenging by the length of the sequences encoding the TET enzymes, which are >12 Kb in length. However, a similar strategy could be employed, whereby the TET enzyme of interest is introduced as a knock-in under an albumin promoter and is, therefore, only expressed as the cells mature and acquire a hepatocyte-like phenotype. An alternative to genetic manipulation is the use of small molecule inhibitors of TET function. One compound, Bobcat339, was recently described and shown to inhibit TET1 and TET2 *in vitro* [426]. Although a more straightforward approach than genetic manipulation, the off-target effects of Bobcat339 have not been described, and it is unclear whether it also inhibits activity of other dioxygenase enzymes.

The use of the monolayer population of HLCs to model NAFLD offers the opportunity to study mitochondrial dysfunction and lipid accumulation in the absence of the confounding effects created by other cell types. Taking this forward, the next logical step would be to use hPSC-based spheroids, which are then differentiated to HLCs within a 3D structure [427]. Compared with the monolayer-based model used here, these spheroids display a more organised architecture and make it easier to extrapolate findings to *in vivo* organs. Furthermore, these spheroids maintain function for up to a year, making them suitable for much longer-term studies than monolayer HLCs [428]. They can also be grafted into mice, where they support liver function of the animal. This offers the opportunity to continue studying a single cell population but *in vivo*, where the cells will receive endocrine signals and interact with other cell populations. Further to this is the option of using liver organoids, which offers the opportunity to study the impact of cellular interactions on the pathology of NAFLD. As well as hepatocyte-like cells, these organoids also contain stellate-like cells and Kupffer-like cell [429]. Recently it was shown that where the iPSCs used to generate these organoids were derived from patients with NASH that the organoids developed a NASH-like phenotype, and transcriptomes that were distinct from those derived from healthy donors [430]. As with spheroid models, it is also possible to graft liver organoids into rodents, to assess their function *in vivo*. These organoid models offer the opportunity to further study genetic susceptibility to progressing from benign steatosis to NASH, as well as opening the door to personalised medicine.

Whilst there is still work to be done, hPSC-derived liver tissues have the potential to be used in clinical settings in the future. Firstly, hPSCs can be differentiated to HLCs in 96-well plates, as demonstrated in this thesis, which opens the door to high throughput screenings of potential therapeutic compounds to treat liver disease. These screenings can be used to measure the impact of these compounds on disease pathology as well as toxicity. Building upon this, once putative therapeutic compounds have been assessed, they can be tested *in vivo*, through grafting of human hepatocyte spheroids into mice [428]. This will enable closer approximations of the effects of these compounds in a clinical setting. Furthermore, drugs such as metformin, appear to exert a beneficial effect on NAFLD [431], but in-depth studies are required to determine the mechanisms underpinning these effects. Using *in vitro* systems such

as the one utilised in this thesis may enable more in-depth mechanistic studies of drugs such as these, which, in turn, may enable the development of more targeted therapeutics. With further development of iPSC-based hepatocytes, this may also enable the development of personalised drug therapies, whereby clinicians administer drugs that patients are most likely to respond favourably to.

6.5 Conclusions

In conclusion, this thesis provides insights into the mitochondrial dysfunction that occurs in association with hepatic steatosis and how rewiring of respiratory pathways induces macrovesicular steatosis. These findings open the door to understanding precisely how increased fumarate synthesis induces intracellular lipid accumulation. Although manipulation of fumarate levels is an unlikely therapeutic strategy, exploration of downstream pathways affected by its accumulation may reveal therapeutic targets. Until a strategy is identified to reverse the accelerating global rates of obesity, targeting NAFLD directly, and preventing disease progression will limit the burden on healthcare systems.

Chapter 7: Bibliography

1. Di Cesare, M. *et al.* Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *The Lancet* vol. 387 1377–1396 (2016).
2. WHO | 10 facts on obesity. *WHO* (2017).
3. Guh, D. P. *et al.* The incidence of co-morbidities related to obesity and overweight: A systematic review and meta-analysis. *BMC Public Health* **9**, 88 (2009).
4. Abd El-Kader, S. M. & El-Den Ashmawy, E. M. S. Non-alcoholic fatty liver disease: The diagnosis and management. *World J. Hepatol.* **7**, 846–858 (2015).
5. Maruhama, Y. *et al.* Interactions of obesity and glucose-stimulated insulin secretion in familial hypertriglyceridemia. *Diabetes* **27**, 682–693 (1978).
6. Adler, M. & Schaffner, F. Fatty liver hepatitis and cirrhosis in obese patients. *Am. J. Med.* **67**, 811–816 (1979).
7. Ludwig, J., Viggiano, T. R., McGill, D. B. & Ott, B. J. Nonalcoholic steatohepatitis. Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clin. Proc.* **55**, 434–438 (1980).
8. Keeffe, E. B., Adesman, P. W., Stenzel, P. & Palmer, R. M. Steatosis and cirrhosis in an obese diabetic - Resolution of fatty liver by fasting. *Dig. Dis. Sci.* **32**, 441–445 (1987).
9. Lee, R. G. Nonalcoholic steatohepatitis: A study of 49 patients. *Hum. Pathol.* **20**, 594–598 (1989).
10. Pennisi, G., Celsa, C., Giammanco, A., Spatola, F. & Petta, S. The burden of hepatocellular carcinoma in non-alcoholic fatty liver disease: Screening issue and future perspectives. *International Journal of Molecular Sciences* vol. 20 (2019).
11. Romero-Gómez, M., Zelber-Sagi, S. & Trenell, M. Treatment of NAFLD with diet, physical activity and exercise. *Journal of Hepatology* vol. 67 829–846 (2017).
12. Younossi, Z. M. *et al.* Global epidemiology of nonalcoholic fatty liver disease—Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* **64**, 73–84 (2016).
13. Marengo, A., Jouness, R. I. K. & Bugianesi, E. Progression and Natural History of Nonalcoholic Fatty Liver Disease in Adults. *Clinics in Liver Disease* vol. 20 313–324 (2016).
14. Ratziu, V., Goodman, Z. & Sanyal, A. Current efforts and trends in the treatment of NASH. *Journal of Hepatology* vol. 62 S65–S75 (2015).
15. Takahashi, Y. & Fukusato, T. Histopathology of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. *World J. Gastroenterol.* **20**, 15539–15548 (2014).
16. Brunt, E. M. & Tiniakos, D. G. Histopathology of nonalcoholic fatty liver disease. *World Journal of Gastroenterology* vol. 16 5286–5296 (2010).
17. Valenti, L., Bugianesi, E., Pajvani, U. & Targher, G. Nonalcoholic fatty liver disease: cause or consequence of type 2 diabetes? *Liver International* vol. 36 1563–1579 (2016).

18. Wei, Y., Rector, R. S., Thyfault, J. P. & Ibdah, J. A. Nonalcoholic fatty liver disease and mitochondrial dysfunction. *World Journal of Gastroenterology* vol. 14 193–199 (2008).
19. Gorden, D. L. *et al.* Biomarkers of NAFLD progression: A lipidomics approach to an epidemic 1. *J. Lipid Res.* **56**, 722–736 (2015).
20. Schwimmer, J. B. *et al.* Histopathology of pediatric nonalcoholic fatty liver disease. *Hepatology* **42**, 641–649 (2005).
21. Puri, P. & Sanyal, A. J. Nonalcoholic fatty liver disease: Definitions, risk factors, and workup. *Clin. Liver Dis.* **1**, 99–103 (2012).
22. George, J. *et al.* NASH and insulin resistance: Insulin hypersecretion and specific association with the insulin resistance syndrome. *Hepatology* **35**, 373–379 (2002).
23. Caldwell, S. *et al.* Hepatocellular ballooning in NASH. *J. Hepatol.* **53**, 719–723 (2010).
24. Duarte, N. *et al.* How inflammation impinges on NAFLD: A role for Kupffer cells. *BioMed Research International* vol. 2015 (2015).
25. Brown, G. T. & Kleiner, D. E. Histopathology of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Metabolism: Clinical and Experimental* vol. 65 1080–1086 (2016).
26. Le, T. H. *et al.* The Zonal Distribution of Megamitochondria with Crystalline Inclusions in Nonalcoholic Steatohepatitis. *Hepatology* **39**, 1423–1429 (2004).
27. Kamath, P. S. & Kim, W. R. The Model for End-stage Liver Disease (MELD). *Hepatology* vol. 45 797–805 (2007).
28. Buzzetti, E., Pinzani, M. & Tsochatzis, E. A. The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD). *Metabolism.* **65**, 1038–1048 (2016).
29. Day, C. P. & James, O. F. W. Steatohepatitis: A tale of two 'Hits'? *Gastroenterology* vol. 114 842–845 (1998).
30. Fang, Y. L., Chen, H., Wang, C. L. & Liang, L. Pathogenesis of non-alcoholic fatty liver disease in children and adolescence: From 'two hit theory' to 'multiple hit model'. *World Journal of Gastroenterology* vol. 24 2974–2983 (2018).
31. Thyfault, J. P. *et al.* Rats selectively bred for low aerobic capacity have reduced hepatic mitochondrial oxidative capacity and susceptibility to hepatic steatosis and injury. *J. Physiol.* **587**, 1805–1816 (2009).
32. Rector, R. S. *et al.* Mitochondrial dysfunction precedes insulin resistance and hepatic steatosis and contributes to the natural history of non-alcoholic fatty liver disease in an obese rodent model. *J. Hepatol.* **52**, 727–736 (2010).
33. Cuyàs, E. *et al.* Metformin regulates global DNA methylation via mitochondrial one-carbon metabolism. *Oncogene* **37**, 963–970 (2018).
34. Sunny, N. E., Parks, E. J., Browning, J. D. & Burgess, S. C. Excessive hepatic mitochondrial TCA cycle and gluconeogenesis in humans with nonalcoholic fatty liver disease. *Cell Metab.* **14**, 804–810 (2011).
35. Nagao, H. *et al.* Increased dynamics of tricarboxylic acid cycle and glutamate synthesis in obese adipose tissue: In vivo metabolic turnover analysis. *J. Biol. Chem.* **292**, 4469–4483 (2017).
36. Neuschwander-Tetri, B. A. Nontriglyceride hepatic lipotoxicity: The new

paradigm for the pathogenesis of NASH. *Current Gastroenterology Reports* vol. 12 49–56 (2010).
37. Yamaguchi, K. *et al.* Inhibiting triglyceride synthesis improves hepatic steatosis but exacerbates liver damage and fibrosis in obese mice with nonalcoholic steatohepatitis. *Hepatology* **45**, 1366–1374 (2007).
38. Subramanian, S. *et al.* Dietary cholesterol exacerbates hepatic steatosis and inflammation in obese LDL receptor-deficient mice. *J. Lipid Res.* **52**, 1626–1635 (2011).
39. Nakamuta, M. *et al.* Impact of cholesterol metabolism and the LXRα-SREBP-1c pathway on nonalcoholic fatty liver disease. *Int. J. Mol. Med.* **23**, 603–608 (2009).
40. Pagadala, M., Kasumov, T., McCullough, A. J., Zein, N. N. & Kirwan, J. P. Role of ceramides in nonalcoholic fatty liver disease. *Trends Endocrinol. Metab.* **23**, 365–371 (2012).
41. Assies, J. *et al.* Effects of oxidative stress on fatty acid- and one-carbon-metabolism in psychiatric and cardiovascular disease comorbidity. *Acta psychiatrica Scandinavica* vol. 130 163–180 (2014).
42. Lockman, K. A. *et al.* Proteomic profiling of cellular steatosis with concomitant oxidative stress in vitro. *Lipids Health Dis.* **15**, 114 (2016).
43. Zhou, H. & Liu, R. ER stress and hepatic lipid metabolism. *Frontiers in Genetics* vol. 5 112 (2014).
44. Liu, X. & Green, R. M. Endoplasmic reticulum stress and liver diseases. *Liver Research* vol. 3 55–64 (2019).
45. Maiers, J. L. & Malhi, H. Endoplasmic Reticulum Stress in Metabolic Liver Diseases and Hepatic Fibrosis. *Semin. Liver Dis.* **39**, 235–248 (2019).
46. Malhotra, J. D. & Kaufman, R. J. ER stress and Its functional link to mitochondria: Role in cell survival and death. *Cold Spring Harb. Perspect. Biol.* **3**, 1–13 (2011).
47. Epstein, S. *et al.* Activation of the unfolded protein response pathway causes ceramide accumulation in yeast and INS-1E insulinoma cells. *J. Lipid Res.* **53**, 412–420 (2012).
48. Luukkonen, P. K. *et al.* Hepatic ceramides dissociate steatosis and insulin resistance in patients with non-alcoholic fatty liver disease. *J. Hepatol.* **64**, 1167–1175 (2016).
49. Narayanan, S., Surette, F. A. & Hahn, Y. S. The immune landscape in nonalcoholic steatohepatitis. *Immune Network* vol. 16 147–158 (2016).
50. Jia, L. *et al.* Hepatocyte toll-like receptor 4 regulates obesity-induced inflammation and insulin resistance. *Nat. Commun.* **5**, 3878 (2014).
51. Kahraman, A. *et al.* Major histocompatibility complex class I-related chains A and B (MIC A/B): A novel role in nonalcoholic steatohepatitis. *Hepatology* **51**, 92–102 (2010).
52. Tsuchida, T. & Friedman, S. L. Mechanisms of hepatic stellate cell activation. *Nature Reviews Gastroenterology and Hepatology* vol. 14 397–411 (2017).
53. Puche, J. E., Saiman, Y. & Friedman, S. L. Hepatic stellate cells and liver fibrosis. *Compr. Physiol.* **3**, 1473–1492 (2013).
54. Flier, J. S., Underhill, L. H. & Friedman, S. L. The Cellular Basis of Hepatic Fibrosis – Mechanisms and Treatment Strategies. *New England Journal of*

Medicine vol. 328 1828–1835 (1993).
55. Friedman, S. L. Hepatic stellate cells: Protean, multifunctional, and enigmatic cells of the liver. *Physiological Reviews* vol. 88 125–172 (2008).
56. Dewidar, Meyer, Dooley & Meindl-Beinker. TGF-β in Hepatic Stellate Cell Activation and Liver Fibrogenesis. *Cells* **8**, 1419 (2019).
57. Angulo, P. *et al.* Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* **149**, 389-397.e10 (2015).
58. Basaranoglu, M., Basaranoglu, G. & Sentürk, H. From fatty liver to fibrosis: A tale of 'second hit'. *World J. Gastroenterol.* **19**, 1158–1165 (2013).
59. Kassi, E., Pervanidou, P., Kaltsas, G. & Chrousos, G. Metabolic syndrome: Definitions and controversies. *BMC Med.* **9**, 48 (2011).
60. NEEL, J. V. Diabetes mellitus: a 'thrifty' genotype rendered detrimental by 'progress'? *Am. J. Hum. Genet.* **14**, 353–362 (1962).
61. Speakman, J. R. Thrifty genes for obesity and the metabolic syndrome - Time to call off the search? *Diabetes and Vascular Disease Research* vol. 3 7–11 (2006).
62. Wood, K. L., Miller, M. H. & Dillon, J. F. Systematic review of genetic association studies involving histologically confirmed non-Alcoholic fatty liver disease. *BMJ Open Gastroenterology* vol. 2 e000019 (2015).
63. Namjou, B. *et al.* GWAS and enrichment analyses of non-alcoholic fatty liver disease identify new trait-associated genes and pathways across eMERGE Network. *BMC Med.* **17**, 135 (2019).
64. Romeo, S. *et al.* Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat. Genet.* **40**, 1461–1465 (2008).
65. Di Costanzo, A. *et al.* Evaluation of Polygenic Determinants of Non-Alcoholic Fatty Liver Disease (NAFLD) By a Candidate Genes Resequencing Strategy. *Sci. Rep.* **8**, 3702 (2018).
66. Pirazzi, C. *et al.* Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro. *J. Hepatol.* **57**, 1276–1282 (2012).
67. Speliotes, E. K. *et al.* Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet.* **7**, e1001324 (2011).
68. Kantartzis, K. *et al.* Dissociation between fatty liver and insulin resistance in humans carrying a variant of the patatin-like phospholipase 3 gene. *Diabetes* **58**, 2616–2623 (2009).
69. Huang, Y., Cohen, J. C. & Hobbs, H. H. Expression and characterization of a PNPLA3 protein isoform (I148M) associated with nonalcoholic fatty liver disease. *J. Biol. Chem.* **286**, 37085–37093 (2011).
70. Chamoun, Z., Vacca, F., Parton, R. G. & Gruenberg, J. PNPLA3/adiponutrin functions in lipid droplet formation. *Biol. Cell* **105**, 219–233 (2013).
71. He, S. *et al.* A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. *J. Biol. Chem.* **285**, 6706–6715 (2010).
72. Corbin, K. D. *et al.* Genetic signatures in choline and 1-carbon metabolism are associated with the severity of hepatic steatosis. *FASEB J.* **27**, 1674–

1689 (2013).
73. Radziejewska, A., Muzsik, A., Milagro, F. I., Martínez, J. A. & Chmurzynska, A. One-Carbon Metabolism and Nonalcoholic Fatty Liver Disease: The Crosstalk between Nutrients, Microbiota, and Genetics. *Lifestyle Genomics* vol. 13 53–63 (2020).
74. Anstee, Q. M. *et al.* Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically-characterised cohort. *J. Hepatol.* **0**, (2020).
75. Manco, M. Insulin Resistance and NAFLD: A Dangerous Liaison beyond the Genetics. *Children* **4**, 74 (2017).
76. Nielsen, S., Guo, Z. K., Johnson, C. M., Hensrud, D. D. & Jensen, M. D. Splanchnic lipolysis in human obesity. *J. Clin. Invest.* **113**, 1582–1588 (2004).
77. Kamagate, A. *et al.* FoxO1 links hepatic insulin action to endoplasmic reticulum stress. *Endocrinology* **151**, 3521–3535 (2010).
78. Sutherland, C., Leighton, I. A. & Cohen, P. Inactivation of glycogen synthase kinase-3β by phosphorylation: New kinase connections in insulin and growth-factor signalling. *Biochem. J.* **296**, 15–19 (1993).
79. Lambert, J. E., Ramos-Roman, M. A., Browning, J. D. & Parks, E. J. Increased de novo lipogenesis is a distinct characteristic of individuals with nonalcoholic fatty liver disease. *Gastroenterology* **146**, 726–735 (2014).
80. Liu, X. & Yao, Z. Chronic over-nutrition and dysregulation of GSK3 in diseases. *Nutrition and Metabolism* vol. 13 49 (2016).
81. Chen, H. *et al.* PI3K-resistant GSK3 controls adiponectin formation and protects from metabolic syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5754–5759 (2016).
82. Fabbrini, E. *et al.* Alterations in adipose tissue and hepatic lipid kinetics in obese men and women with nonalcoholic fatty liver disease. *Gastroenterology* **134**, 424–431 (2008).
83. Donnelly, K. L. *et al.* Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *J. Clin. Invest.* **115**, 1343–1351 (2005).
84. McGarry, J. D. & Foster, D. W. Regulation of hepatic fatty acid oxidation and ketone body production. *Annu. Rev. Biochem.* **49**, 395–420 (1980).
85. Ipsen, D. H., Lykkesfeldt, J. & Tveden-Nyborg, P. Molecular mechanisms of hepatic lipid accumulation in non-alcoholic fatty liver disease. *Cellular and Molecular Life Sciences* vol. 75 3313–3327 (2018).
86. Holt, H. B. *et al.* Non-esterified fatty acid concentrations are independently associated with hepatic steatosis in obese subjects. *Diabetologia* **49**, 141–148 (2006).
87. Greco, D. *et al.* Gene expression in human NAFLD. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **294**, G1281-7 (2008).
88. Miquilena-Colina, M. E. *et al.* Hepatic fatty acid translocase CD36 upregulation is associated with insulin resistance, hyperinsulinaemia and increased steatosis in non-alcoholic steatohepatitis and chronic hepatitis C. *Gut* **60**, 1394–1402 (2011).
89. Targher, G. *et al.* Associations between plasma adiponectin concentrations and liver histology in patients with nonalcoholic fatty liver disease. *Clin.*

Endocrinol. (Oxf). **64**, 679–683 (2006).
90. Buechler, C., Wanninger, J. & Neumeier, M. Adiponectin, a key adipokine in obesity related liver diseases. *World Journal of Gastroenterology* vol. 17 2801–2811 (2011).
91. Yamauchi, T. & Kadowaki, T. Physiological and pathophysiological roles of adiponectin and adiponectin receptors in the integrated regulation of metabolic and cardiovascular diseases. *International Journal of Obesity* vol. 32 S13–S18 (2008).
92. Lehrskov, L. L. & Christensen, R. H. The role of interleukin-6 in glucose homeostasis and lipid metabolism. *Seminars in Immunopathology* vol. 41 491–499 (2019).
93. Wallenius, V. *et al.* Interleukin-6-deficient mice develop mature-onset obesity. *Nat. Med.* **8**, 75–79 (2002).
94. Matthews, V. B. *et al.* Interleukin-6-deficient mice develop hepatic inflammation and systemic insulin resistance. *Diabetologia* **53**, 2431–2441 (2010).
95. Paredes-Turrubiarte, G. *et al.* Severity of non-alcoholic fatty liver disease is associated with high systemic levels of tumor necrosis factor alpha and low serum interleukin 10 in morbidly obese patients. *Clin. Exp. Med.* **16**, 193–202 (2016).
96. Wandrer, F. *et al.* TNF-Receptor-1 inhibition reduces liver steatosis, hepatocellular injury and fibrosis in NAFLD mice. *Cell Death Dis.* **11**, 212 (2020).
97. Bi, J. *et al.* PPARγ alleviated hepatocyte steatosis through reducing SOCS3 by inhibiting JAK2/STAT3 pathway. *Biochem. Biophys. Res. Commun.* **498**, 1037–1044 (2018).
98. Peng, K. Y. *et al.* Mitochondrial dysfunction-related lipid changes occur in nonalcoholic fatty liver disease progression. *J. Lipid Res.* **59**, 1977–1986 (2018).
99. Ameer, F., Scandiuzzi, L., Hasnain, S., Kalbacher, H. & Zaidi, N. De novo lipogenesis in health and disease. *Metabolism: Clinical and Experimental* vol. 63 895–902 (2014).
100. Ruderman, N. B., Saha, A. K. & Kraegen, E. W. Malonyl CoA, AMP-Activated Protein Kinase, and Adiposity. *Endocrinology* vol. 144 5166–5171 (2003).
101. Ortega-Prieto, P. & Postic, C. Carbohydrate sensing through the transcription factor ChREBP. *Front. Genet.* **10**, 472 (2019).
102. Dif, N. *et al.* Insulin activates human sterol-regulatory-element-binding protein-1c (SREBP-1c) promoter through SRE motifs. *Biochem. J.* **400**, 179–188 (2006).
103. Benhamed, F. *et al.* The lipogenic transcription factor ChREBP dissociates hepatic steatosis from insulin resistance in mice and humans. *J. Clin. Invest.* **122**, 2176–2194 (2012).
104. Knebel, B. *et al.* Liver-specific expression of transcriptionally active srebp-1c is associated with fatty liver and increased visceral fat mass. *PLoS One* **7**, e31812 (2012).
105. Higuchi, N. *et al.* Liver X receptor in cooperation with SREBP-1c is a major lipid synthesis regulator in nonalcoholic fatty liver disease. *Hepatol. Res.* **38**,

1122–1129 (2008).
106. Feingold, K. R. & Grunfeld, C. *Introduction to Lipids and Lipoproteins*. *Endotext* (MDText.com, Inc., 2000).
107. Tiwari, S. & Siddiqi, S. A. Intracellular trafficking and secretion of VLDL. *Arteriosclerosis, Thrombosis, and Vascular Biology* vol. 32 1079–1086 (2012).
108. Ding, W. X. & Yin, X. M. Mitophagy: Mechanisms, pathophysiological roles, and analysis. *Biological Chemistry* vol. 393 547–564 (2012).
109. Satapati, S. *et al.* Elevated TCA cycle function in the pathology of diet-induced hepatic insulin resistance and fatty liver. *J. Lipid Res.* **53**, 1080–1092 (2012).
110. Patterson, R. E. *et al.* Lipotoxicity in steatohepatitis occurs despite an increase in tricarboxylic acid cycle activity. *Am. J. Physiol. - Endocrinol. Metab.* **310**, ajpendo.00492.2015 (2016).
111. Sunny, N. E., Bril, F. & Cusi, K. Mitochondrial adaptation in nonalcoholic fatty liver disease: novel mechanisms and treatment strategies. *Trends Endocrinol. Metab.* **28**, 250–260 (2017).
112. Sanyal, A. J. *et al.* Nonalcoholic steatohepatitis: association of insulin resistance and mitochondrial abnormalities. *Gastroenterology* **120**, 1183–1192 (2001).
113. Caldwell, S. H. *et al.* Mitochondrial abnormalities in non-alcoholic steatohepatitis. *J. Hepatol.* **31**, 430–434 (1999).
114. Zhang, J. Autophagy and mitophagy in cellular damage control. *Redox Biology* vol. 1 19–23 (2013).
115. Kim, Y. *et al.* PINK1 controls mitochondrial localization of Parkin through direct phosphorylation. *Biochem. Biophys. Res. Commun.* **377**, 975–980 (2008).
116. Jin, S. M. *et al.* Mitochondrial membrane potential regulates PINK1 import and proteolytic destabilization by PARL. *J. Cell Biol.* **191**, 933–942 (2010).
117. Edmunds, L. R. *et al.* Liver-specific Prkn knockout mice are more susceptible to diet-induced hepatic steatosis and insulin resistance. *Mol. Metab.* 101051 (2020) doi:10.1016/J.MOLMET.2020.101051.
118. Akram, M. Citric acid cycle and role of its intermediates in metabolism. *Cell Biochemistry and Biophysics* vol. 68 475–478 (2014).
119. McCommis, K. S. & Finck, B. N. Mitochondrial pyruvate transport: a historical perspective and future research directions. *Biochem. J.* **466**, 443–454 (2015).
120. Owen, O. E., Kalhan, S. C. & Hanson, R. W. The key role of anaplerosis and cataplerosis for citric acid cycle function. *Journal of Biological Chemistry* vol. 277 30409–30412 (2002).
121. Tannahill, G. M. *et al.* Succinate is an inflammatory signal that induces IL-1β through HIF-1α. *Nature* **496**, 238–242 (2013).
122. Raimundo, N., Baysal, B. E. & Shadel, G. S. Revisiting the TCA cycle: signaling to tumor formation. *Trends Mol. Med.* **17**, 641–649 (2011).
123. Nagaraj, R. *et al.* Nuclear localization of mitochondrial TCA cycle enzymes as a critical step in mammalian zygotic genome activation. *Cell* **168**, 210-223.e11 (2017).
124. Sinton, M. C., Hay, D. C. & Drake, A. J. Metabolic control of gene

transcription in non-alcoholic fatty liver disease: the role of the epigenome. *Clinical epigenetics* vol. 11 104 (2019).

125. Brekke, E. *et al.* Direct measurement of backflux between oxaloacetate and fumarate following pyruvate carboxylation. *Glia* **60**, 147–158 (2012).
126. Sutherland, G., Tyson, R. & Auer, R. Truncation of the Krebs cycle during hypoglycemic coma. *Med. Chem. (Los. Angeles).* **4**, 379–385 (2008).
127. Tyrakis, P. A. *et al.* Fumarate hydratase loss causes combined respiratory chain defects. *Cell Rep.* **21**, 1036–1047 (2017).
128. Lyall, M. J. *et al.* Modelling non-alcoholic fatty liver disease in human hepatocyte-like cells. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, 20170362 (2018).
129. Satapati, S. *et al.* Mitochondrial metabolism mediates oxidative stress and inflammation in fatty liver. *J. Clin. Invest.* **125**, 4447–4462 (2015).
130. Koliaki, C. *et al.* Adaptation of hepatic mitochondrial function in humans with non-alcoholic fatty liver is lost in steatohepatitis. *Cell Metab.* **21**, 739–746 (2015).
131. Knott, A. B., Perkins, G., Schwarzenbacher, R. & Bossy-Wetzel, E. Mitochondrial fragmentation in neurodegeneration. *Nature Reviews Neuroscience* vol. 9 505–518 (2008).
132. McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J. & Schofield, C. J. Structural studies on human 2-oxoglutarate dependent oxygenases. *Curr. Opin. Struct. Biol.* **20**, 659–672 (2010).
133. Laukka, T. *et al.* Fumarate and succinate regulate expression of hypoxia-inducible genes via TET enzymes. *J. Biol. Chem.* **291**, 4256–4265 (2016).
134. Roach, P. L. *et al.* Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. *Nature* **375**, 700–704 (1995).
135. Hausinger, R. P. Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes. *Critical Reviews in Biochemistry and Molecular Biology* vol. 39 21–68 (2004).
136. Xiao, M. *et al.* Inhibition of α-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev.* **26**, 1326–1338 (2012).
137. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science (80-.).* **324**, 930–5 (2009).
138. Pollard, P. J. *et al.* Regulation of Jumonji-domain-containing histone demethylases by hypoxia-inducible factor (HIF)-1α. *Biochem. J.* **416**, 387–394 (2008).
139. Stiehl, D. P. *et al.* Increased prolyl 4-hydroxylase domain proteins compensate for decreased oxygen levels: Evidence for an autoregulatory oxygen-sensing system. *J. Biol. Chem.* **281**, 23482–23491 (2006).
140. Qi, H. H. *et al.* Prolyl 4-hydroxylation regulates Argonaute 2 stability. *Nature* **455**, 421–424 (2008).
141. Hutton, J. J., Kaplan, A. & Udenfriend, S. Conversion of the amino acid sequence Gly-Pro-Pro in protein to Gly-Pro-Hyp by collagen proline hydroxylase. *Arch. Biochem. Biophys.* **121**, 384–391 (1967).
142. Bhatnagar, R. S., Rapaka, R. S. & Urry, D. W. Interaction of polypeptide

models of elastin with prolyl hydroxylase. *FEBS Lett.* **95**, 61–64 (1978).
143. Ke, Q. & Costa, M. Hypoxia-inducible factor-1 (HIF-1). *Mol. Pharmacol.* **70**, 1469–1480 (2006).
144. Gorres, K. L. & Raines, R. T. Prolyl 4-hydroxylase. *Critical Reviews in Biochemistry and Molecular Biology* vol. 45 106–124 (2010).
145. Koh, M. Y., Darnay, B. G. & Powis, G. Hypoxia-associated factor, a novel E3-ubiquitin ligase, binds and ubiquitinates hypoxia-inducible factor 1, leading to its oxygen-independent degradation. *Mol. Cell. Biol.* **28**, 7081–7095 (2008).
146. Chua, Y. L. *et al.* Stabilization of hypoxia-inducible factor-1α protein in hypoxia occurs independently of mitochondrial reactive oxygen species production. *J. Biol. Chem.* **285**, 31277–31284 (2010).
147. Koivunen, P. *et al.* Inhibition of hypoxia-inducible factor (HIF) hydroxylases by citric acid cycle intermediates: Possible links between cell metabolism and stabilization of HIF. *J. Biol. Chem.* **282**, 4524–4532 (2007).
148. Furuta, E. *et al.* Fatty acid synthase gene is up-regulated by hypoxia via activation of Akt and sterol regulatory element binding protein-1. *Cancer Res.* **68**, 1003–1011 (2008).
149. Carabelli, J. *et al.* High fat diet-induced liver steatosis promotes an increase in liver mitochondrial biogenesis in response to hypoxia. *J. Cell. Mol. Med.* **15**, 1329–1338 (2011).
150. Kietzmann, T., Cornesse, Y., Brechtel, K., Modaressi, S. & Jungermann, K. Perivenous expression of the mRNA of the three hypoxia-inducible factor alpha-subunits, HIF1alpha, HIF2alpha and HIF3alpha, in rat liver. *Biochem. J.* **354**, 531–7 (2001).
151. Hosogai, N. *et al.* Adipose tissue hypoxia in obesity and its impact on adipocytokine dysregulation. *Diabetes* **56**, 901–911 (2007).
152. Rausch, M. E., Weisberg, S., Vardhana, P. & Tortoriello, D. V. Obesity in C57BL/6J mice is characterized by adipose tissue hypoxia and cytotoxic T-cell infiltration. *Int. J. Obes.* **32**, 451–463 (2008).
153. Mesarwi, O. A. *et al.* Hepatocyte hypoxia inducible factor-1 mediates the development of liver fibrosis in a mouse model of nonalcoholic fatty liver disease. *PLoS One* **11**, e0168572 (2016).
154. Bird, A. Perceptions of epigenetics. *Nature* vol. 447 396–398 (2007).
155. Dupont, C., Armant, D. R. & Brenner, C. A. Epigenetics: Definition, mechanisms and clinical perspective. *Seminars in Reproductive Medicine* vol. 27 351–357 (2009).
156. Berger, S. L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. An operational definition of epigenetics. *Genes Dev.* **23**, 781–783 (2009).
157. Jablonka, E. V. A. & Raz, G. A. L. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q. Rev. Biol.* **84**, 131–176 (2009).
158. Cartier, J. *et al.* Investigation into the role of the germline epigenome in the transmission of glucocorticoid-programmed effects across generations. *Genome Biol.* **19**, 50 (2018).
159. Costa, S. & Shaw, P. 'Open minded' cells: how cells can change fate. *Trends Cell Biol.* **17**, 101–106 (2007).
160. Banno, K. *et al.* Epimutation and cancer: A new carcinogenic mechanism of

Lynch syndrome. *International Journal of Oncology* vol. 41 793–797 (2012).
161. Novak, K. Epigenetics changes in cancer cells. in *MedGenMed Medscape General Medicine* vol. 6 17 (WebMD/Medscape Health Network, 2004).
162. Rothbart, S. B. & Strahl, B. D. Interpreting the language of histone and DNA modifications. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* vol. 1839 627–643 (2014).
163. Mariño-Ramírez, L., Kann, M. G., Shoemaker, B. A. & Landsman, D. Histone structure and nucleosome stability. *Expert Review of Proteomics* vol. 2 719–729 (2005).
164. Burgess, R. J. & Zhang, Z. Histone chaperones in nucleosome assembly and human disease. *Nature Structural and Molecular Biology* vol. 20 14–22 (2013).
165. Avvakumov, N., Nourani, A. & Côté, J. Histone Chaperones: Modulators of Chromatin Marks. *Molecular Cell* vol. 41 502–514 (2011).
166. Narlikar, G. J., Sundaramoorthy, R. & Owen-Hughes, T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* vol. 154 490–503 (2013).
167. Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nature Reviews Molecular Cell Biology* vol. 16 178–189 (2015).
168. Nathan, D., Sterner, D. E. & Berger, S. L. Histone modifications: Now summoning sumoylation. *Proc. Natl. Acad. Sci.* **100**, 13118–13120 (2003).
169. Tamaru, H. Confining euchromatin/heterochromatin territory: Jumonji crosses the line. *Genes and Development* vol. 24 1465–1478 (2010).
170. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science (80-.).* **298**, 1039–1043 (2002).
171. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).
172. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001).
173. Schotta, G. *et al.* A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **18**, 1251–1262 (2004).
174. Kourmouli, N. *et al.* Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. *J. Cell Sci.* **117**, 2491–2501 (2004).
175. Krogan, N. J. *et al.* The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: Linking transcriptional elongation to histone methylation. *Mol. Cell* **11**, 721–729 (2003).
176. Bernstein, B. E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169–181 (2005).
177. Ng, H. H., Robert, F., Young, R. A. & Struhl, K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* **11**, 709–719 (2003).
178. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* vol. 403 41–45 (2000).
179. Turner, B. M. Histone acetylation and an epigenetic code. *BioEssays* **22**,

836–845 (2000).
180. Shogren-Knaak, M. *et al.* Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science (80-.).* **311**, 844–847 (2006).
181. Anand, R. & Marmorstein, R. Structure and mechanism of lysine-specific demethylase enzymes. *Journal of Biological Chemistry* vol. 282 35425–35429 (2007).
182. Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* **7**, 715–727 (2006).
183. Grewal, S. I. & Rice, J. C. Regulation of heterochromatin by histone methylation and small RNAs. *Current Opinion in Cell Biology* vol. 16 230–238 (2004).
184. Cervera, A. M., Bayley, J. P., Devilee, P. & McCreath, K. J. Inhibition of succinate dehydrogenase dysregulates histone modification in mammalian cells. *Mol. Cancer* **8**, 89 (2009).
185. Jones, P. A. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics* vol. 13 484–492 (2012).
186. Duncan, C. G. *et al.* Dosage compensation and DNA methylation landscape of the X chromosome in mouse liver. *Sci. Rep.* **8**, 10138 (2018).
187. Robertson, K. D. DNA methylation and human disease. *Nature Reviews Genetics* vol. 6 597–610 (2005).
188. Stadler, M. B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).
189. Charlet, J. *et al.* Bivalent Regions of Cytosine Methylation and H3K27 Acetylation Suggest an Active Role for DNA Methylation at Enhancers. *Mol. Cell* **62**, 422–431 (2016).
190. Choy, M. K. *et al.* Genome-wide conserved consensus transcription factor binding motifs are hyper-methylated. *BMC Genomics* **11**, 519 (2010).
191. Defossez, P. A. & Stancheva, I. Biological functions of methyl-CpG-binding proteins. in *Progress in Molecular Biology and Translational Science* vol. 101 377–398 (Prog Mol Biol Transl Sci, 2011).
192. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
193. Smith, Z. D. & Meissner, A. DNA methylation: Roles in mammalian development. *Nature Reviews Genetics* vol. 14 204–220 (2013).
194. Suzuki, M. M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465–476 (2008).
195. Leitch, H. G. *et al.* Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20**, 311–316 (2013).
196. Bachman, M. *et al.* 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat. Chem.* **6**, 1049–1055 (2014).
197. Bird, A. The dinucleotide CG as a genomic signalling module. *Journal of Molecular Biology* vol. 409 47–53 (2011).
198. Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S. & Heintz, N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417–1430 (2012).
199. Song, C. X. *et al.* Selective chemical labeling reveals the genome-wide

distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol.* **29**, 68–75 (2011).
200. Ito, S. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science (80-.).* **333**, 1300–1303 (2011).
201. Thomson, J. P. & Meehan, R. R. The application of genome-wide 5-hydroxymethylcytosine studies in cancer research. *Epigenomics* vol. 9 77–91 (2017).
202. Lin, I. H., Chen, Y. F. & Hsu, M. T. Correlated 5-hydroxymethylcytosine (5hmC) and gene expression profiles underpin gene and organ-specific epigenetic regulation in adult mouse brain and liver. *PLoS One* **12**, e0170779 (2017).
203. Putiri, E. L. *et al.* Distinct and overlapping control of 5-methylcytosine and 5-hydroxymethylcytosine by the TET proteins in human cancer cells. *Genome Biol.* **15**, R81 (2014).
204. Wu, H. *et al.* Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.* **25**, 679–684 (2011).
205. Shi, D. Q., Ali, I., Tang, J. & Yang, W. C. New insights into 5hmC DNA modification: Generation, distribution and function. *Frontiers in Genetics* vol. 8 (2017).
206. Szulwach, K. E. *et al.* Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet.* **7**, e1002154 (2011).
207. Tan, L. *et al.* Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. *Nucleic Acids Res.* **41**, e84 (2013).
208. Lyall, M. J. *et al.* Non-alcoholic fatty liver disease (NAFLD) is associated with dynamic changes in DNA hydroxymethylation. *Epigenetics* **15**, 61–71 (2020).
209. Rose, N. R., McDonough, M. A., King, O. N. F., Kawamura, A. & Schofield, C. J. Inhibition of 2-oxoglutarate dependent oxygenases. *Chemical Society Reviews* vol. 40 4364–4397 (2011).
210. He, Y. F. *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science (80-.).* **333**, 1303–1307 (2011).
211. Ono, R. *et al.* LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res.* **62**, 4075–4080 (2002).
212. Huang, Y. *et al.* Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 1361–1366 (2014).
213. Sowers, L. C. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
214. Lorsbach, R. B. *et al.* TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia* **17**, 637–641 (2003).
215. Nestor, C. E. *et al.* Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res.* **22**, 467–477 (2012).

216. Kriaucionis, S. & Heintz, N. The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science (80-.).* **324**, 929–930 (2009).
217. Wu, Y. *et al.* Oct4 and the small molecule inhibitor, SC1, regulates Tet2 expression in mouse embryonic stem cells. *Mol. Biol. Rep.* **40**, 2897–2906 (2013).
218. Zhao, Z. *et al.* Combined Loss of Tet1 and Tet2 Promotes B Cell, but Not Myeloid Malignancies, in Mice. *Cell Rep.* **13**, 1692–1704 (2015).
219. Shen, L. *et al.* Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell* **15**, 459–471 (2014).
220. Hu, L. *et al.* Structural insight into substrate preference for TET-mediated oxidation. *Nature* **527**, 118–122 (2015).
221. Jin, S. G. *et al.* Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. *Cell Rep.* **14**, 493–505 (2016).
222. Wu, H. *et al.* Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**, 389–393 (2011).
223. Ge, W. *et al.* Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science (80-.).* **329**, 444–448 (2010).
224. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. *Genes Dev.* **16**, 2893–2905 (2002).
225. Chen, Q. R. *et al.* Multi-SNP Analysis of GWAS Data Identifies Pathways Associated with Nonalcoholic Fatty Liver Disease. *PLoS One* **8**, e65982 (2013).
226. Vella, S. *et al.* EZH2 down-regulation exacerbates lipid accumulation and inflammation in in vitro and in vivo NAFLD. *Int. J. Mol. Sci.* **14**, 24154–24168 (2013).
227. Haltiwanger, R. S., Blomberg, M. A. & Hart, G. W. Glycosylation of nuclear and cytoplasmic proteins: Purification and characterization of a uridine diphospho-N-acetylglucosamine:polypeptide ??-N-acetylglucosaminyltransferase. *J. Biol. Chem.* **267**, 9005–9013 (1992).
228. Hrit, J. *et al.* OGT binds a conserved C-terminal domain of TET1 to regulate TET1 activity and function in development. *Elife* **7**, (2018).
229. Kim, Y. H., Nakayama, T. & Nayak, J. Glycolysis and the hexosamine biosynthetic pathway as novel targets for upper and lower airway inflammation. *Allergy, Asthma and Immunology Research* vol. 10 6–11 (2018).
230. Ferrer, C. M. *et al.* O-GlcNAcylation regulates cancer metabolism and survival stress signaling via regulation of the HIF-1 pathway. *Mol. Cell* **54**, 820–831 (2014).
231. Xu, W. *et al.* O-GlcNAc transferase promotes fatty liver-associated liver cancer through inducing palmitic acid and activating endoplasmic reticulum stress. *J. Hepatol.* **67**, 310–320 (2017).
232. Vella, P. *et al.* Tet proteins connect the O-linked N-acetylglucosamine

transferase OGT to chromatin in embryonic stem cells. *Mol. Cell* **49**, 645–656 (2013).
233. Zhang, Y. J. *et al.* Global hypomethylation in hepatocellular carcinoma and its relationship to aflatoxin B1 exposure. *World J. Hepatol.* **4**, 169–175 (2012).
234. Lu, C. & Thompson, C. B. Metabolic regulation of epigenetics. *Cell Metabolism* vol. 16 9–17 (2012).
235. Tateishi, K., Okada, Y., Kallin, E. M. & Zhang, Y. Role of Jhdm2a in regulating metabolic gene expression and obesity resistance. *Nature* **458**, 757–761 (2009).
236. Jun, H. J., Kim, J., Hoang, M. H. & Lee, S. J. Hepatic lipid accumulation alters global histone H3 lysine 9 and 4 trimethylation in the peroxisome proliferator-activated receptor alpha network. *PLoS One* **7**, e44345 (2012).
237. Pogribny, I. P. *et al.* Hepatic epigenetic phenotype predetermines individual susceptibility to hepatic steatosis in mice fed a lipogenic methyl-deficient diet. *J. Hepatol.* **51**, 176–186 (2009).
238. Lyall, M. J. *et al.* Methyl donor deficient diets cause distinct alterations in lipid metabolism but are poorly representative of human NAFLD. *Wellcome Open Res.* **2**, 67 (2017).
239. Pirola, C. J. *et al.* Epigenetic modifications in the biology of nonalcoholic fatty liver disease: The role of DNA hydroxymethylation and TET proteins. *Med. (United States)* **94**, e1480 (2015).
240. Hotta, K. *et al.* Identification of the genomic region under epigenetic regulation during non-alcoholic fatty liver disease progression. *Hepatol. Res.* **48**, E320–E334 (2018).
241. Murphy, S. K. *et al.* Relationship between methylome and transcriptome in patients with nonalcoholic fatty liver disease. *Gastroenterology* **145**, 1076–1087 (2013).
242. Ahrens, M. *et al.* DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. *Cell Metab.* **18**, 296–302 (2013).
243. Hotta, K. *et al.* Identification of differentially methylated region (DMR) networks associated with progression of nonalcoholic fatty liver disease. *Sci. Rep.* **8**, 13567 (2018).
244. de Mello, V. D. *et al.* Human liver epigenetic alterations in non-alcoholic steatohepatitis are related to insulin action. *Epigenetics* **12**, 287–295 (2017).
245. Gerhard, G. S. *et al.* Differentially methylated loci in NAFLD cirrhosis are associated with key signaling pathways. *Clin. Epigenetics* **10**, 93 (2018).
246. Sookoian, S. *et al.* Epigenetic regulation of insulin resistance in nonalcoholic fatty liver disease: Impact of liver methylation of the peroxisome proliferator-activated receptor γ coactivator 1α promoter. *Hepatology* **52**, 1992–2000 (2010).
247. Thomson, J. P. *et al.* Loss of tet1-associated5-hydroxymethylcytosine is concomitant with aberrant promoter hypermethylation in liver cancer. *Cancer Res.* **76**, 3097–3108 (2016).
248. Ito, M. *et al.* Longitudinal analysis of murine steatohepatitis model induced by chronic exposure to high-fat diet. *Hepatol. Res.* **37**, 50–57 (2007).
249. Gallou-Kabani, C. *et al.* C57BL/6J and A/J mice fed a high-fat diet delineate

components of metabolic syndrome. *Obesity* **15**, 1996–2005 (2007).
250. Ishimoto, T. *et al.* High-fat and high-sucrose (western) diet induces steatohepatitis that is dependent on fructokinase. *Hepatology* **58**, 1632–1643 (2013).
251. Rinella, M. E. & Green, R. M. The methionine-choline deficient dietary model of steatohepatitis does not exhibit insulin resistance. *J. Hepatol.* **40**, 47–51 (2004).
252. Anstee, Q. M. & Goldin, R. D. Mouse models in non-alcoholic fatty liver disease and steatohepatitis research. *International Journal of Experimental Pathology* vol. 87 1–16 (2006).
253. Llufrio, E. M., Wang, L., Naser, F. J. & Patti, G. J. Sorting cells alters their redox state and cellular metabolome. *Redox Biol.* **16**, 381–387 (2018).
254. Pavlova, N. N. & Thompson, C. B. The Emerging Hallmarks of Cancer Metabolism. *Cell Metabolism* vol. 23 27–47 (2016).
255. Huang, Q. *et al.* Metabolic characterization of hepatocellular carcinoma using nontargeted tissue metabolomics. *Cancer Res.* **73**, 4992–5002 (2013).
256. Liu, C. *et al.* Decrease of 5-Hydroxymethylcytosine Is Associated with Progression of Hepatocellular Carcinoma through Downregulation of TET1. *PLoS One* **8**, e62828 (2013).
257. Xu, Y. *et al.* Genome-wide Regulation of 5hmC, 5mC, and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells. *Mol. Cell* **42**, 451–464 (2011).
258. Chen, M. L. *et al.* Quantification of 5-methylcytosine and 5-hydroxymethylcytosine in genomic DNA from hepatocellular carcinoma tissues by capillary hydrophilic-interaction liquid chromatography/quadrupole TOF mass spectrometry. *Clin. Chem.* **59**, 824–832 (2013).
259. Hernandez-Vargas, H. *et al.* Hepatocellular Carcinoma Displays Distinct DNA Methylation Signatures with Potential as Clinical Predictors. *PLoS One* **5**, e9749 (2010).
260. Hirsova, P. *et al.* Lipid-Induced Signaling Causes Release of Inflammatory Extracellular Vesicles from Hepatocytes. *Gastroenterology* **150**, 956–967 (2016).
261. Szkolnicka, D. & Hay, D. C. Concise Review: Advances in Generating Hepatocytes from Pluripotent Stem Cells for Translational Medicine. *Stem Cells* vol. 34 1421–1426 (2016).
262. Medine, C. N. *et al.* Developing High-Fidelity Hepatotoxicity Models From Pluripotent Stem Cells. *Stem Cells Transl. Med.* **2**, 505–509 (2013).
263. Sun, P., Zhou, X., Farnworth, S. L., Patel, A. H. & Hay, D. C. Modeling human liver biology using stem cell-derived hepatocytes. *International Journal of Molecular Sciences* vol. 14 22011–22021 (2013).
264. Sullivan, G. J. *et al.* Generation of functional human hepatic endoderm from human induced pluripotent stem cells. *Hepatology* **51**, 329–335 (2010).
265. Doss, M. X. & Sachinidis, A. Current Challenges of iPSC-Based Disease Modeling and Therapeutic Implications. *Cells* **8**, 403 (2019).
266. Ohnuki, M. & Takahashi, K. Present and future challenges of induced pluripotent stem cells. *Philosophical Transactions of the Royal Society B: Biological Sciences* vol. 370 20140367 (2015).

267. Wang, Y. *et al.* Defined and scalable generation of hepatocyte-like cells from human pluripotent stem cells. *J. Vis. Exp.* e55355–e55355 (2017).
268. Greally, J. M. & Drake, A. J. The current state of epigenetic research in humans. *JAMA Pediatr.* **171**, 103 (2016).
269. Huang, Y. *et al.* The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* **5**, e8888 (2010).
270. Short, K. R. *et al.* Decline in skeletal muscle mitochondrial function with aging in humans. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 5618–5623 (2005).
271. Boutagy, N. E. *et al.* Isolation of mitochondria from minimal quantities of mouse skeletal muscle for high throughput microplate respiratory measurements. *J. Vis. Exp.* **2015**, (2015).
272. Orellana, E. & Kasinski, A. Sulforhodamine B (SRB) assay in cell culture to investigate cell proliferation. *Bio-Protocol* **6**, (2016).
273. Hollinshead, K. E. R. *et al.* Oncogenic IDH1 Mutations Promote Enhanced Proline Synthesis through PYCR1 to Support the Maintenance of Mitochondrial Redox Homeostasis. *Cell Rep.* **22**, 3107–3114 (2018).
274. Ludwig, C. & Günther, U. L. MetaboLab - advanced NMR data processing and analysis for metabolomics. *BMC Bioinformatics* **12**, 366 (2011).
275. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
276. Kazimierczuk, K. & Orekhov, V. Y. Accelerated NMR spectroscopy by using compressed sensing. *Angew. Chemie - Int. Ed.* **50**, 5556–5559 (2011).
277. Orekhov, V. Y. & Jaravine, V. A. Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Progress in Nuclear Magnetic Resonance Spectroscopy* vol. 59 271–292 (2011).
278. Smith, S. A., Levante, T. O., Meier, B. H. & Ernst, R. R. Computer Simulations in Magnetic Resonance. An Object-Oriented Programming Approach. *J. Magn. Reson. Ser. A* **106**, 75–105 (1994).
279. Wishart, D. S. *et al.* HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Res.* **41**, D801-7 (2013).
280. Battello, N. *et al.* The role of HIF-1 in oncostatin M-dependent metabolic reprogramming of hepatic cells. *Cancer Metab.* **4**, 3 (2016).
281. Hiller, K. *et al.* Metabolite detector: Comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis. *Anal. Chem.* **81**, 3429–3439 (2009).
282. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
283. Andrews, S. FastQC: a quality control tool for high throughput sequence data. *2010* http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
284. Dobin, A. & Gingeras, T. R. Mapping RNA-seq Reads with STAR. *Curr. Protoc. Bioinforma.* **51**, 11.14.1-11.14.19 (2015).
285. Picard Tools - By Broad Institute. http://broadinstitute.github.io/picard/ (2018).
286. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
287. Babicki, S. *et al.* Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* **44**, W147–W153 (2016).

288. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
289. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **47**, D590–D595 (2019).
290. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).
291. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
292. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
293. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
294. Stark, R. & Brown, G. Differential Binding Analysis of ChIP-Seq Peak Data. http://bioconductor.org/packages/release/bioc/html/DiffBind.html.
295. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).
296. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187-91 (2014).
297. Wang, L. & Yu, S. Pathology of non-alcoholic fatty liver disease. *Int. J. Dig. Dis.* **2**, (2016).
298. Suppli, M. P. *et al.* Hepatic transcriptome signatures in patients with varying degrees of non-alcoholic fatty liver disease compared to healthy normal-weight individuals. *Am. J. Physiol. Gastrointest. Liver Physiol.* **316**, G462–G472 (2019).
299. Egnatchik, R. A. *et al.* Glutamate– oxaloacetate transaminase activity promotes palmitate lipotoxicity in rat hepatocytes by enhancing anaplerosis and citric acid cycle flux. *J. Biol. Chem.* **294**, 3081–3090 (2019).
300. Knobeloch, D. *et al.* Human hepatocytes: Isolation, culture, and quality procedures. *Methods Mol. Biol.* **806**, 99–120 (2012).
301. Godoy, P. *et al.* Gene network activity in cultivated primary hepatocytes is highly similar to diseased mammalian liver tissue. *Arch. Toxicol.* **90**, 2513–2529 (2016).
302. Bray, M. A. *et al.* Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat. Protoc.* **11**, 1757–1774 (2016).
303. Sinton, M. *et al.* Nonalcoholic fatty liver disease is associated with decreased hepatocyte mitochondrial respiration, but not mitochondrial number. *bioRxiv* 2020.03.10.985200 (2020) doi:10.1101/2020.03.10.985200.
304. Sinton, M. C. *et al.* Macrovesicular steatosis in nonalcoholic fatty liver disease is a consequence of purine nucleotide cycle driven fumarate accumulation. *bioRxiv* 2020.06.23.166728 (2020) doi:10.1101/2020.06.23.166728.
305. Wobser, H. *et al.* Lipid accumulation in hepatocytes induces fibrogenic

activation of hepatic stellate cells. *Cell Res.* **19**, 996–1005 (2009).
306. Wanninger, J. *et al.* Lipid accumulation impairs adiponectin-mediated induction of activin A by increasing TGFbeta in primary human hepatocytes. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1811**, 626–633 (2011).
307. Sahini, N. & Borlak, J. Genomics of human fatty liver disease reveal mechanistically linked lipid droplet–associated gene regulations in bland steatosis and nonalcoholic steatohepatitis. *Transl. Res.* **177**, 41–69 (2016).
308. Fujii, H. *et al.* Expression of perilipin and adipophilin in nonalcoholic fatty liver disease; relevance to oxidative injury and hepatocyte ballooning. *J. Atheroscler. Thromb.* **16**, 893–901 (2009).
309. Chen, W. *et al.* Inactivation of Plin4 downregulates Plin5 and reduces cardiac lipid accumulation in mice. *Am. J. Physiol. - Endocrinol. Metab.* **304**, E770-9 (2013).
310. Wang, C. *et al.* Perilipin 5 improves hepatic lipotoxicity by inhibiting lipolysis. *Hepatology* **61**, 870–882 (2015).
311. Kerkar, N. Non-alcoholic steatohepatitis in children. in *Pediatric Transplantation* vol. 8 613–618 (John Wiley & Sons, Ltd, 2004).
312. Collison, K. S. *et al.* Effect of dietary monosodium glutamate on trans fat-induced nonalcoholic fatty liver disease. *J. Lipid Res.* **50**, 1521–1537 (2009).
313. Kolwankar, D. *et al.* Association Between Nonalcoholic Hepatic Steatosis and Hepatic Cytochrome P-450 3A Activity. *Clin. Gastroenterol. Hepatol.* **5**, 388–393 (2007).
314. Liu, H., Pathak, P., Boehme, S. & Chiang, J. Y. L. Cholesterol 7α-hydroxylase protects the liver from inflammation and fibrosis by maintaining cholesterol homeostasis. *J. Lipid Res.* **57**, 1831–1844 (2016).
315. Nikolaou, N. *et al.* AKR1D1 is a novel regulator of metabolic phenotype in human hepatocytes and is dysregulated in non-alcoholic fatty liver disease. *Metabolism.* **99**, 67–80 (2019).
316. Hardwick, R. N. *et al.* Altered UDP-glucuronosyltransferase and sulfotransferase expression and function during progressive stages of human nonalcoholic fatty liver diseases. *Drug Metab. Dispos.* **41**, 554–561 (2013).
317. Schiöth, H. B. *et al.* A targeted analysis reveals relevant shifts in the methylation and transcription of genes responsible for bile acid homeostasis and drug metabolism in non-alcoholic fatty liver disease. *BMC Genomics* **17**, 462 (2016).
318. Yamaguchi, M. & Murata, T. Involvement of regucalcin in lipid metabolism and diabetes. *Metabolism: Clinical and Experimental* vol. 62 1045–1051 (2013).
319. Hart-Unger, S. *et al.* Hormone signaling and fatty liver in females: Analysis of estrogen receptor α mutant mice. *Int. J. Obes.* **41**, 945–954 (2017).
320. Saloniemi, T., Jokela, H., Strauss, L., Pakarinen, P. & Poutanen, M. The diversity of sex steroid action: Novel functions of hydroxysteroid (17β) dehydrogenases as revealed by genetically modified mouse models. *Journal of Endocrinology* vol. 212 27–40 (2012).
321. Chiang, J. Y. L. Targeting bile acids and lipotoxicity for NASH treatment. *Hepatol. Commun.* **1**, 1002–1004 (2017).
322. Pullinger, C. R. *et al.* Human cholesterol 7α-hydroxylase (CYP7A1) deficiency

has a hypercholesterolemic phenotype. *J. Clin. Invest.* **110**, 109–117 (2002).
323. Woolsey, S. J., Mansell, S. E., Kim, R. B., Tirona, R. G. & Beaton, M. D. CYP3A activity and expression in nonalcoholic fatty liver disease. *Drug Metab. Dispos.* **43**, 1484–1490 (2015).
324. Yalcin, E. B. *et al.* Downregulation of sulfotransferase expression and activity in diseased human livers. *Drug Metab. Dispos.* **41**, 1642–1650 (2013).
325. Yamaguchi, T. *et al.* Characterization of lipid droplets in steroidogenic MLTC-1 Leydig cells: Protein profiles and the morphological change induced by hormone stimulation. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1851**, 1285–1295 (2015).
326. Wyrick, J. J. *et al.* Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* **402**, 418–421 (1999).
327. Sancho, M., Diani, E., Beato, M. & Jordan, A. Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth. *PLoS Genet.* **4**, e1000227 (2008).
328. Pazienza, V. *et al.* SIRT1-metabolite binding histone macroH2A1.1 protects hepatocytes against lipid accumulation. *Aging (Albany. NY).* **6**, 35–47 (2014).
329. Pagani, M. *et al.* Endoplasmic reticulum oxidoreductin 1-Lβ (ERO1-Lβ), a human gene induced in the course of the unfolded protein response. *J. Biol. Chem.* **275**, 23685–23692 (2000).
330. Van Galen, P. *et al.* The unfolded protein response governs integrity of the haematopoietic stem-cell pool during stress. *Nature* **510**, 268–272 (2014).
331. Puri, P. *et al.* Activation and Dysregulation of the Unfolded Protein Response in Nonalcoholic Fatty Liver Disease. *Gastroenterology* **134**, 568–576 (2008).
332. Rinella, M. E. *et al.* Dysregulation of the unfolded protein response in db/db mice with diet-induced steatohepatitis. *Hepatology* **54**, 1600–1609 (2011).
333. Chen, Z., Tian, R., She, Z., Cai, J. & Li, H. Role of oxidative stress in the pathogenesis of nonalcoholic fatty liver disease. *Free Radical Biology and Medicine* vol. 152 116–141 (2020).
334. Vercesi, A. E. *et al.* Mitochondrial calcium transport and the redox nature of the calcium-induced membrane permeability transition. *Free Radical Biology and Medicine* vol. 129 1–24 (2018).
335. García-Ruiz, I. *et al.* Mitochondrial complex i subunits are decreased in murine nonalcoholic fatty liver disease: Implication of peroxynitrite. *J. Proteome Res.* **9**, 2450–2459 (2010).
336. Mantena, S. K. *et al.* High fat diet induces dysregulation of hepatic oxygen gradients and mitochondrial function in vivo. *Biochem. J.* **417**, 183–193 (2009).
337. Pérez-Carreras, M. *et al.* Defective hepatic mitochondrial respiratory chain in patients with nonalcoholic steatohepatitis. *Hepatology* **38**, 999–1007 (2003).
338. Godoy, P. *et al.* Gene networks and transcription factor motifs defining the differentiation of stem cells into hepatocyte-like cells. *J. Hepatol.* **63**, 934–942 (2015).
339. Rauckhorst, A. J. *et al.* The mitochondrial pyruvate carrier mediates high fat diet-induced increases in hepatic TCA cycle capacity. *Mol. Metab.* **6**, 1468–1479 (2017).
340. Antoniewicz, M. R. A guide to 13C metabolic flux analysis for the cancer

biologist. *Experimental and Molecular Medicine* vol. 50 19 (2018).
341. Wang, Y. *et al.* Multiomics Analyses of HNF4α Protein Domain Function during Human Pluripotent Stem Cell Differentiation. *iScience* **16**, 206–217 (2019).
342. Scaini, G. *et al.* Toxicity of octanoate and decanoate in rat peripheral tissues: evidence of bioenergetic dysfunction and oxidative damage induction in liver and skeletal muscle. *Mol. Cell. Biochem.* **361**, 329–335 (2012).
343. Wu, Z. *et al.* Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**, 115–124 (1999).
344. Hui, S. *et al.* Glucose feeds the TCA cycle via circulating lactate. *Nature* **551**, 115–118 (2017).
345. Ray, I., Dasgupta, A. & De, R. K. Succinate aggravates NAFLD progression to liver cancer on the onset of obesity: An in silico model. *J. Bioinform. Comput. Biol.* **16**, 1850008 (2018).
346. Lardy, H. A., Paetkau, V. & Walter, P. Paths of carbon in gluconeogenesis and lipogenesis: the role of mitochondria in supplying precursors of phosphoenolpyruvate. *Proc. Natl. Acad. Sci. U. S. A.* **53**, 1410–1415 (1965).
347. Fletcher, J. A. *et al.* Impaired ketogenesis and increased acetyl-CoA oxidation promote hyperglycemia in human fatty liver. *JCI Insight* **4**, (2019).
348. Gastaldelli, A. *et al.* Relationship Between Hepatic/Visceral Fat and Hepatic Insulin Resistance in Nondiabetic and Type 2 Diabetic Subjects. *Gastroenterology* **133**, 496–506 (2007).
349. Das, K. *et al.* Nonobese population in a developing country has a high prevalence of nonalcoholic fatty liver and significant liver disease. *Hepatology* **51**, 1593–1602 (2010).
350. Singh, S. P. *et al.* Nonalcoholic fatty liver disease (NAFLD) without insulin resistance: Is it different? *Clin. Res. Hepatol. Gastroenterol.* **39**, 482–488 (2015).
351. Phipps, M. & Wattacheril, J. Non-alcoholic fatty liver disease (NAFLD) in non-obese individuals. *Frontline Gastroenterology* flgastro-2018-101119 (2019) doi:10.1136/flgastro-2018-101119.
352. Dasika, S. K., Vinnakota, K. C. & Beard, D. A. Determination of the catalytic mechanism for mitochondrial malate dehydrogenase. *Biophys. J.* **108**, 408–419 (2015).
353. Perry, R. J. *et al.* Propionate increases hepatic pyruvate cycling and anaplerosis and alters mitochondrial metabolism. *J. Biol. Chem.* **291**, 12161–12170 (2016).
354. Perry, R. J., Peng, L., Cline, G. W., Petersen, K. F. & Shulman, G. I. A Non-invasive Method to Assess Hepatic Acetyl-CoA In Vivo. *Cell Metab.* **25**, 749–756 (2017).
355. Goedeke, L. *et al.* Acetyl-CoA Carboxylase Inhibition Reverses NAFLD and Hepatic Insulin Resistance but Promotes Hypertriglyceridemia in Rodents. *Hepatology* **68**, 2197–2211 (2018).
356. Li, Q., Zhang, S., Berthiaume, J. M., Simons, B. & Zhang, G. F. Novel approach in LC-MS/MS using MRM to generate a full profile of acyl-CoAs: Discovery of acyl-dephospho-CoAs. *J. Lipid Res.* **55**, 592–602 (2014).
357. Lynn, R. & Guynn, R. W. Equilibrium constants under physiological conditions

for the reactions of succinyl coenzyme A synthetase and the hydrolysis of succinyl coenzyme A to coenzyme A and succinate. *J. Biol. Chem.* **253**, 2546–53 (1978).

358. Chouchani, E. T. *et al.* Ischaemic accumulation of succinate controls reperfusion injury through mitochondrial ROS. *Nature* **515**, 431–435 (2014).
359. Ajalla Aleixo, M. A., Rangel, V. L., Rustiguel, J. K., de Pádua, R. A. P. & Nonato, M. C. Structural, biochemical and biophysical characterization of recombinant human fumarate hydratase. *FEBS J.* **286**, 1925–1940 (2019).
360. Porcelli, A. M. *et al.* The genetic and metabolic signature of oncocytic transformation implicates HIF1α destabilization. *Hum. Mol. Genet.* **19**, 1019–1032 (2010).
361. Arinze, I. J. Facilitating understanding of the purine nucleotide cycle and the one-carbon pool: Part I: The purine nucleotide cycle. *Biochemistry and Molecular Biology Education* vol. 33 165–168 (2005).
362. De Chiara, F. *et al.* Ammonia Scavenging Prevents Progression of Fibrosis in Experimental Nonalcoholic Fatty Liver Disease. *Hepatology* **71**, 874–892 (2020).
363. Sangineto, M. *et al.* Dimethyl fumarate ameliorates hepatic inflammation in alcohol related liver disease. *Liver Int.* liv.14483 (2020) doi:10.1111/liv.14483.
364. Rensen, S. S. *et al.* Activation of the complement system in human nonalcoholic fatty liver disease. *Hepatology* **50**, 1809–1817 (2009).
365. Handa, P. *et al.* Differences in hepatic expression of iron, inflammation and stress-related genes in patients with nonalcoholic steatohepatitis. *Ann. Hepatol.* **16**, 77–85 (2017).
366. La Bonte, L. R., Dokken, B., Davis-Gorman, G., Stahl, G. L. & McDonagh, P. F. The mannose-binding lectin pathway is a significant contributor to reperfusion injury in the type 2 diabetic heart. *Diabetes Vasc. Dis. Res.* **6**, 172–180 (2009).
367. Castellano-Castillo, D. *et al.* Complement factor C3 methylation and mRNA expression is associated to BMI and insulin resistance in obesity. *Genes (Basel).* **9**, (2018).
368. Cai, G., Huang, Z., Yu, L. & Li, L. A preliminary study showing no association between methylation levels of C3 gene promoter and the risk of CAD. *Lipids Health Dis.* **18**, 5 (2019).
369. Fargion, S. *et al.* Hyperferritinemia, iron overload, and multiple metabolic alterations identify patients at risk for nonalcoholic steatohepatitis. *Am. J. Gastroenterol.* **96**, 2448–2455 (2001).
370. George, D. K. *et al.* Increased hepatic iron concentration in nonalcoholic steatohepatitis is associated with increased fibrosis. *Gastroenterology* **114**, 311–318 (1998).
371. Senates, E. *et al.* Serum levels of hepcidin in patients with biopsy-proven nonalcoholic fatty liver disease. *Metab. Syndr. Relat. Disord.* **9**, 287–290 (2011).
372. Auguet, T. *et al.* Hepcidin in morbidly obese women with nonalcoholic fatty liver disease. *PLoS One* **12**, (2017).
373. Huang, H., Taraboletti, A. & Shriver, L. P. Dimethyl fumarate modulates antioxidant and lipid metabolism in oligodendrocytes. *Redox Biol.* **5**, 169–175

(2015).
374. Bhargava, P. *et al.* Dimethyl fumarate treatment induces lipid metabolism alterations that are linked to immunological changes. *Ann. Clin. Transl. Neurol.* **6**, 33–45 (2019).
375. Rappez, L. *et al.* Spatial single-cell profiling of intracellular metabolomes in situ. *bioRxiv* 510222 (2019) doi:10.1101/510222.
376. Zhao, P. & Saltiel, A. R. From overnutrition to liver injury: AMP-activated protein kinase in nonalcoholic fatty liver diseases. *The Journal of biological chemistry* vol. 295 12279–12289 (2020).
377. Alkhouri, N., Lawitz, E., Noureddin, M., DeFronzo, R. & Shulman, G. I. GS-0976 (Firsocostat): an investigational liver-directed acetyl-CoA carboxylase (ACC) inhibitor for the treatment of non-alcoholic steatohepatitis (NASH). *Expert Opin. Investig. Drugs* **29**, 135–141 (2020).
378. Boudaba, N. *et al.* AMPK Re-Activation Suppresses Hepatic Steatosis but its Downregulation Does Not Promote Fatty Liver Development. *EBioMedicine* **28**, 194–209 (2018).
379. Rabinovitch, R. C. *et al.* AMPK Maintains Cellular Metabolic Homeostasis through Regulation of Mitochondrial Reactive Oxygen Species. *Cell Rep.* **21**, 1–9 (2017).
380. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
381. Guo, W. *et al.* 3D RNA-seq - a powerful and flexible tool for rapid and accurate differential expression and alternative splicing analysis of RNA-seq data for biologists. *bioRxiv* 656686 (2019) doi:10.1101/656686.
382. Wei, S. *et al.* Identification of hub genes and construction of transcriptional regulatory network for the progression of colon adenocarcinoma hub genes and TF regulatory network of colon adenocarcinoma. *J. Cell. Physiol.* **235**, 2037–2048 (2020).
383. Zhang, M. *et al.* The Landscape of DNA Methylation Associated With the Transcriptomic Network of Intramuscular Adipocytes Generates Insight Into Intramuscular Fat Deposition in Chicken. *Front. Cell Dev. Biol.* **8**, (2020).
384. Parafati, M., Kirby, R. J., Khorasanizadeh, S., Rastinejad, F. & Malany, S. A nonalcoholic fatty liver disease model in human induced pluripotent stem cell-derived hepatocytes, created by endoplasmic reticulum stress-induced steatosis. *DMM Dis. Model. Mech.* **11**, (2018).
385. Atanasovska, B. *et al.* A liver-specific long noncoding RNA with a role in cell viability is elevated in human nonalcoholic steatohepatitis. *Hepatology* **66**, 794–808 (2017).
386. Revill, K. *et al.* Genome-wide methylation analysis and epigenetic unmasking identify tumor suppressor genes in hepatocellular carcinoma. *Gastroenterology* **145**, 1424-35.e1–25 (2013).
387. Cazanave, S. *et al.* The Transcriptomic Signature of Disease Development and Progression of Nonalcoholic Fatty Liver Disease. *Sci. Rep.* **7**, 17193 (2017).
388. Wilson, C. H. & Kumar, S. Caspases in metabolic disease and their therapeutic potential. *Cell Death and Differentiation* vol. 25 1010–1024 (2018).

389. Marina, R. J. *et al.* TET -catalyzed oxidation of intragenic 5-methylcytosine regulates CTCF -dependent alternative splicing. *EMBO J.* **35**, 335–355 (2016).
390. Villena, J. A. New insights into PGC-1 coactivators: Redefining their role in the regulation of mitochondrial function and beyond. *FEBS Journal* vol. 282 647–672 (2015).
391. Rodríguez-Aguilera, J. R. *et al.* Genome-wide 5-hydroxymethylcytosine (5hmC) emerges at early stage of in vitro differentiation of a putative hepatocyte progenitor. *Sci. Rep.* **10**, 7822 (2020).
392. Tabish, A. M. *et al.* Association of intronic DNA methylation and hydroxymethylation alterations in the epigenetic etiology of dilated cardiomyopathy. *Am. J. Physiol. - Hear. Circ. Physiol.* **317**, H168–H180 (2019).
393. Nestor, C. E. *et al.* 5-Hydroxymethylcytosine Remodeling Precedes Lineage Specification during Differentiation of Human CD4+ T Cells. *Cell Rep.* **16**, 559–570 (2016).
394. Wehner, D. *et al.* Wnt signaling controls pro-regenerative Collagen XII in functional spinal cord regeneration in zebrafish. *Nat. Commun.* **8**, 126 (2017).
395. Yuan, J. *et al.* CUL4B activates Wnt/β-catenin signalling in hepatocellular carcinoma by repressing Wnt antagonists. *J. Pathol.* **235**, 784–795 (2015).
396. Wong, C. M., Fan, S. T. & Ng, I. O. L. β-catenin mutation and overexpression in hepatocellular carcinoma: Clinicopathologic and prognostic significance. *Cancer* **92**, 136–145 (2001).
397. Tian, Y. *et al.* Histone deacetylase HDAC8 promotes insulin resistance and β-catenin activation in NAFLD-associated hepatocellular carcinoma. *Cancer Res.* **75**, 4803–4816 (2015).
398. Spruijt, C. G. *et al.* Dynamic readers for 5-(Hydroxy)methylcytosine and its oxidized derivatives. *Cell* **152**, 1146–1159 (2013).
399. Jones, P. L. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **19**, 187–191 (1998).
400. Fuks, F. *et al.* The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J. Biol. Chem.* **278**, 4035–4040 (2003).
401. Baiocchini, A. *et al.* Extracellular matrix molecular remodeling in human liver fibrosis evolution. *PLoS One* **11**, e0151736 (2016).
402. Cavalcante, R. G., Patil, S., Park, Y., Rozek, L. S. & Sartor, M. A. Integrating DNA methylation and hydroxymethylation data with the mint pipeline. *Cancer Res.* **77**, e27–e30 (2017).
403. Tu, B. P. & Weissman, J. S. Oxidative protein folding in eukaryotes: Mechanisms and consequences. *Journal of Cell Biology* vol. 164 341–346 (2004).
404. Awazawa, M. *et al.* Deregulation of Pancreas-Specific Oxidoreductin ERO1 in the Pathogenesis of Diabetes Mellitus. *Mol. Cell. Biol.* **34**, 1290–1299 (2014).
405. Swanson, K. V., Deng, M. & Ting, J. P.-Y. The NLRP3 inflammasome: molecular activation and regulation to therapeutics. *Nature Reviews Immunology* vol. 19 477–489 (2019).
406. Menu, P. *et al.* ER stress activates the NLRP3 inflammasome via an UPR-independent pathway. *Cell Death Dis.* **3**, e261 (2012).

407. Kanda, T. *et al.* Apoptosis and non-alcoholic fatty liver diseases. *World Journal of Gastroenterology* vol. 24 2661–2672 (2018).
408. Hye Khan, M. A., Schmidt, J., Stavniichuk, A., Imig, J. D. & Merk, D. A dual farnesoid X receptor/soluble epoxide hydrolase modulator treats non-alcoholic steatohepatitis in mice. *Biochem. Pharmacol.* **166**, 212–221 (2019).
409. Cai, S. P. *et al.* Transmembrane protein 88 attenuates liver fibrosis by promoting apoptosis and reversion of activated hepatic stellate cells. *Mol. Immunol.* **80**, 58–67 (2016).
410. Ma, R. *et al.* Promoter methylation of Wnt/β-Catenin signal inhibitor TMEM88 is associated with unfavorable prognosis of nonsmall cell lung cancer. *Cancer Biol. Med.* **14**, 377–386 (2017).
411. Lochmann, T. L., Thomas, R. R., Bennett, J. P. & Taylor, S. M. Epigenetic modifications of the PGC-1α promoter during exercise induced expression in mice. *PLoS One* **10**, e0129647 (2015).
412. Miura, S., Kai, Y., Kamei, Y. & Ezaki, O. Isoform-specific increases in murine skeletal muscle peroxisome proliferator-activated receptor-γ coactivator-1α (PGC-1α) mRNA in response to β2-adrenergic receptor activation and exercise. *Endocrinology* **149**, 4527–4533 (2008).
413. Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S. & Jacobsen, S. E. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.* **12**, R54 (2011).
414. Hu, H. *et al.* Epigenomic landscape of 5-hydroxymethylcytosine reveals its transcriptional regulation of lncRNAs in colorectal cancer. *Br. J. Cancer* **116**, 658–668 (2017).
415. Wang, P., Yan, Y., Yu, W. & Zhang, H. Role of ten-eleven translocation proteins and 5-hydroxymethylcytosine in hepatocellular carcinoma. *Cell Proliferation* vol. 52 (2019).
416. Booth, M. J. *et al.* Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. *Nat. Protoc.* **8**, 1841–1851 (2013).
417. Han, J., He, Y., Zhao, H. & Xu, X. Hypoxia inducible factor-1 promotes liver fibrosis in nonalcoholic fatty liver disease by activating PTEN/p65 signaling pathway. *J. Cell. Biochem.* **120**, 14735–14744 (2019).
418. Kim, J. H., Jung, D. Y., Nagappan, A. & Jung, M. H. Histone H3K9 demethylase JMJD2B induces hepatic steatosis through upregulation of PPARγ2. *Sci. Rep.* **8**, (2018).
419. Nguyen-Lefebvre, A. T. & Horuzsko, A. Kupffer Cell Metabolism and Function. *J. Enzymol. Metab.* **1**, (2015).
420. Picard, M., Taivassalo, T., Gouspillou, G. & Hepple, R. T. Mitochondria: Isolation, structure and function. *Journal of Physiology* vol. 589 4413–4421 (2011).
421. Cappel, D. A. *et al.* Pyruvate-Carboxylase-Mediated Anaplerosis Promotes Antioxidant Capacity by Sustaining TCA Cycle and Redox Metabolism in Liver. *Cell Metab.* **29**, 1291-1305.e8 (2019).
422. Thomson, J. P. *et al.* Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. *Genome Biol.* **13**, R93 (2012).
423. Thomson, J. P. *et al.* Dynamic changes in 5-hydroxymethylation signatures underpin early and late events in drug exposed liver. *Nucleic Acids Res.* **41**,

5639–5654 (2013).
424. Bedogni, G. *et al.* Relationship between glucose metabolism and non-alcoholic fatty liver disease severity in morbidly obese women. *J. Endocrinol. Invest.* **37**, 739–744 (2014).
425. Zhong, X. *et al.* Ten-eleven translocation-2 (Tet2) is involved in myogenic differentiation of skeletal myoblast cells in vitro. *Sci. Rep.* **7**, 43539 (2017).
426. Chua, G. N. L. *et al.* Cytosine-Based TET Enzyme Inhibitors. *ACS Med. Chem. Lett.* **10**, 180–185 (2019).
427. Lucendo-Villarin, B. *et al.* Serum free production of three-dimensional human hepatospheres from pluripotent stem cells. *J. Vis. Exp.* **2019**, e59965 (2019).
428. Rashidi, H. *et al.* 3D human liver tissue from pluripotent stem cells displays stable phenotype in vitro and supports compromised liver function in vivo. *Arch. Toxicol.* **92**, 3117–3129 (2018).
429. Ouchi, R. *et al.* Modeling Steatohepatitis in Humans with Pluripotent Stem Cell-Derived Organoids. *Cell Metab.* **30**, 374-384.e6 (2019).
430. Mccarron, S. *et al.* NASH patient liver derived organoids exhibit patient specific NASH phenotypes and drug responses. *bioRxiv* 791467 (2019) doi:10.1101/791467.
431. Yan, Li., Liu, L., Wang, B., Wang, J. & Chen, D. Metformin in non-alcoholic fatty liver disease: A systematic review and meta-analysis. *Biomed. Reports* **1**, 57–64 (2013).

Chapter 8: Appendices

8.1 Appendix 1: Up and downregulated KEGG pathways

Pathways with upregulated genes		
Pathway	*Fold enrichment*	*Genes*
hsa04713: Circadian entrainment	4.20	ADCY4, ADCY2, CACNA1I, GRIN1, GRIA3, PRKG1, PRKCB, KCNJ5, PLCB4, GRIN2D, CACNA1G, RYR1, GNG2, PER3, GNG4, MTNR1A
hsa04724: Glutamatergic synapse	4.15	ADCY4, DLGAP1, ADCY2, GRIK2, GRIN1, GRIK5, GRIN3B, GRIA3, SHANK1, PRKCB, SLC17A7, GLS2, GRM4, PLCB4, GRIN2D, PLA2G4F, GNG2, GNG4, PLA2G4D
hsa04725: Cholinergic synapse	3.59	ADCY4, ACHE, ADCY2, KCNJ12, KCNJ14, PRKCB, KCNJ4, KCNQ4, KCNQ3, CHRM4, PLCB4, GNG2, PIK3R5, GNG4, KCNQ1, CHRNA3
hsa04750: Inflammatory mediator regulation of TRP channels	3.31	ADCY4, ADCY2, TRPM8, TRPV2, ASIC3, ASIC1, PRKCB, PLCB4, PLA2G4F, PIK3R5, ALOX12, NGF, PLA2G4D
hsa05414: Dilated cardiomyopathy	3.26	ADCY4, ADCY2, ADRB1, SGCG, CACNG8, ITGA8, ITGA7, CACNB1, CACNB4, TNNI3, ITGA2B
hsa04911: Insulin secretion	3.23	TRPM4, ADCY4, ADCY2, PLCB4, KCNN1, KCNN3, ATP1A3, RIMS2, PCLO, KCNJ11, PRKCB
hsa04726: Serotonergic synapse	3.14	GABRB2, PRKCB, KCNJ5, PLCB4, ALOX15B, HTR7, SLC18A2, PLA2G4F, GNG2, HTR1D, HTR3A, GNG4, ALOX12, PLA2G4D
hsa04514: Cell adhesion molecules (CAMs)	2.81	ICAM1, NTNG1, NTNG2, CLDN10, CD40, HLA-DMB, CDH4, HLA-G, NCAM2, ITGA8, NLGN4X, PECAM1, CNTN1, HLA-DOA, CD6, ICOSLG
hsa04921: Oxytocin signalling pathway	2.66	ADCY4, ADCY2, CACNG8, CACNB1, CACNB4, KCNJ12, TRPM2, KCNJ14, PRKCB, KCNJ5, KCNJ4, PLCB4, RYR1, PLA2G4F, NFATC1, PLA2G4D

hsa04080: Neuroactive ligand-receptor interaction	2.43	F2RL2, C5AR1, GABRB2, GRIK2, GRIN1, GABBR1, GRIK5, LPAR3, GRIA3, GRIN3B, P2RX5, GRM4, SSTR2, ADRB1, S1PR1, CHRM4, CHRNA9, SSTR1, GRIN2D, HTR7, CHRNA5, ADRA1A, CALCRL, HTR1D, CHRNA3, GRID1, MTNR1A
hsa04024: cAMP signalling pathway	2.39	ADCY4, HCN2, ADCY2, GRIN1, GABBR1, ATP1A3, GRIN3B, GRIA3, TNNI3, GLI1, SSTR2, ADRB1, SSTR1, PDE4A, GRIN2D, PIK3R5, HTR1D, HCAR2, NFATC1
hsa04020: Calcium signalling pathway	2.37	SLC8A3, ADCY4, SLC8A2, ADCY2, CACNA1I, GRIN1, PRKCB, P2RX5, PLCB4, ADRB1, ATP2A3, GRIN2D, HTR7, CACNA1G, RYR1, ADRA1A, PLCD1
hsa04014: Ras signalling pathway	2.32	FGF19, FGF5, FLT4, FGF17, EFNA3, GRIN1, FGF11, PRKCB, RASAL1, HTR7, RASGRP2, PLA2G4F, PLA1A, GNG2, PIK3R5, NGFR, SYNGAP1, GNG4, RASA4, PLA2G4D, NGF
Pathways with downregulated genes		
Pathway	*Fold enrichment*	*Genes*
hsa05322: Systemic lupus erythematosus	6.65	HIST1H2AB, C7, HIST1H4L, HIST1H2AG, C6, HIST1H2AE, HIST1H2BO, HIST2H2AB, HIST1H2BM, HIST1H4A, HIST1H2BL, HIST1H2BI, HIST2H2AC, HIST1H2BJ, H2AFX, HIST3H2BB, HIST1H4I, HIST1H4J, HIST2H3A, HIST1H3J, HIST1H2BB, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2BH, ACTN2, HIST2H3C, HIST2H3D, HIST2H2BF, HIST1H3A, HIST1H3B, HIST1H2AI, HIST1H2AH, HIST1H3C, HIST1H3E, HIST1H2AJ, HIST1H3F, HIST1H2AM, HIST1H3G, HIST1H2AL, HIST1H3H, HIST1H3I
hsa00140: Steroid hormone biosynthesis	6.43	HSD3B2, CYP3A4, CYP3A5, CYP3A7, HSD3B1, CYP11A1, HSD17B1, UGT1A9, CYP7A1, UGT2B11, UGT2B4, HSD17B6, UGT2B10,

		UGT2A3, SULT1E1, UGT2B15, AKR1D1, CYP19A1
hsa00053: Ascorbate and aldarate metabolism	5.37	UGT1A9, UGT2B11, RGN, UGT2B4, UGT2B10, UGT2A3, UGT2B15
hsa04610: Complement and coagulation cascades	5.10	KNG1, F11, PLAT, MBL2, C7, MASP2, F13A1, C6, F9, C4BPB, C4BPA, F13B, F3, KLKB1, SERPINA5, SERPIND1, CPB2
hsa00040: Pentose and glucuronate interconversions	5.02	UGT1A9, KL, AKR1B10, UGT2B11, UGT2B4, UGT2B10, UGT2A3, UGT2B15
hsa05034: Alcoholism	4.92	HIST1H2AB, HIST1H4L, HIST1H2AG, HIST1H2AE, HIST1H2BO, HIST2H2AB, HIST1H2BM, HIST1H4A, HIST1H2BL, HIST1H2BI, HIST2H2AC, HIST1H2BJ, H2AFX, HIST3H2BB, HIST1H4I, HIST1H4J, HIST2H3A, HIST1H3J, HIST1H2BB, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2BH, FOSB, HIST2H3C, HIST2H3D, HIST2H2BF, NTRK2, HIST1H3A, HIST1H3B, HIST1H2AI, HIST1H2AH, HIST1H3C, HIST1H3E, HIST1H2AJ, HIST1H3F, HIST1H2AM, HIST1H3G, HIST1H2AL, HIST1H3H, HIST1H3I
hsa00982: Drug metabolism - cytochrome P450	4.88	GSTA1, CYP3A4, GSTA2, CYP3A5, ALDH3B2, ADH1B, ADH1A, FMO5, UGT1A9, FMO1, ADH4, UGT2B11, UGT2B4, UGT2A3, UGT2B10, UGT2B15
hsa00830: Retinol metabolism	4.86	CYP3A4, CYP3A5, UGT1A9, CYP3A7, ADH4, UGT2B11, ADH1B, UGT2B4, CYP26A1, HSD17B6, ADH1A, UGT2B10, UGT2A3, UGT2B15, RDH5
hsa05204: Chemical carcinogenesis	4.40	GSTA1, CYP3A4, GSTA2, CYP3A5, CYP3A7, NAT2, ADH1B, ALDH3B2, ADH1A, CYP3A43, UGT1A9, ADH4, UGT2B11, UGT2B4, UGT2A3, UGT2B10, UGT2B15
hsa00983: Drug metabolism - other enzymes	4.05	CYP3A4, UGT1A9, NAT2, UGT2B11, UGT2B4, UGT2B10, UGT2A3, UGT2B15, TK1

hsa00980: Metabolism of xenobiotics by cytochrome P450	3.92	GSTA1, CYP3A4, CYP3A5, GSTA2, ALDH3B2, ADH1B, ADH1A, UGT1A9, ADH4, UGT2B11, UGT2B4, UGT2A3, UGT2B10, UGT2B15
hsa04110: Cell cycle	3.01	CDC6, CDK1, PKMYT1, TTK, CDC20, ESPL1, PTTG1, MCM2, CDC25C, CCNB1, CDKN1C, MAD2L1, CCNB2, PLK1, CDKN2C, BUB1, BUB1B, CCNA2
hsa04114: Oocyte meiosis	2.61	CCNB1, CDK1, MAD2L1, CCNB2, PLK1, SGO1, BUB1, FBXO43, PKMYT1, AURKA, ESPL1, CDC20, PTTG1, CDC25C
hsa05202: Transcriptional misregulation in cancer	2.36	PLAT, HIST2H3A, NFKBIZ, HIST1H3J, MMP9, MMP3, HIST2H3C, HIST2H3D, HHEX, EYA1, CDKN2C, HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I

8.2 Appendix 2: Dysregulated genes in the TCA Cycle KEGG pathway

Ensembl ID	*Gene*	*Log_2FC*	*padj*
ENSG00000197444	OGDHL	1.08547769	3.29E-30
ENSG00000166411	IDH3A	0.74709247	7.70E-20
ENSG00000204370	SDHD	0.3777972	0.0003
ENSG00000100412	ACO2	0.26021698	0.002
ENSG00000146701	MDH2	0.23945724	0.0009
ENSG00000073578	SDHA	-0.1593416	0.03
ENSG00000138413	IDH1	-0.1812103	0.03
ENSG00000168291	PDHB	-0.2026402	0.02
ENSG00000131828	PDHA1	-0.2107741	0.01
ENSG00000163541	SUCLG1	-0.2111892	0.006
ENSG00000143252	SDHC	-0.2414483	0.02
ENSG00000150768	DLAT	-0.2619025	0.0003
ENSG00000091483	FH	-0.4936542	1.01E-12
ENSG00000014641	MDH1	-0.5085681	3.36E-13
ENSG00000101365	IDH3B	-0.5405112	5.07E-13
ENSG00000067829	IDH3G	-0.5829934	1.14E-09
ENSG00000100889	PCK2	-0.5852916	3.07E-18
ENSG00000131473	ACLY	-0.6074454	9.80E-24
ENSG00000173599	PC	-0.9409933	3.13E-37
ENSG00000182054	IDH2	-1.0677314	8.52E-45

8.3 Appendix 3: Dysregulated genes in the Oxidative Phosphorylation KEGG pathway

Ensembl ID	*Gene*	*Log_2FC*	*padj*
ENSG00000147614	ATP6V0D2	3.973933876	9.35E-140
ENSG00000198763	ND2	1.473526224	2.53E-06
ENSG00000198888	ND1	1.363322749	6.19E-05
ENSG00000198840	ND3	1.217622162	2.54E-33
ENSG00000212907	ND4L	1.095495883	0.0001
ENSG00000198886	ND4	0.866690501	0.002
ENSG00000114573	ATP6V1A	0.803262336	2.80E-32
ENSG00000198804	COX1	0.656245362	1.17E-09
ENSG00000198899	ATP6	0.634543196	0.009
ENSG00000198938	COX3	0.632275024	2.80E-10
ENSG00000171130	ATP6V0E2	0.589272395	2.52E-11
ENSG00000198727	CYTB	0.557687429	0.02
ENSG00000047249	ATP6V1H	0.506014583	5.94E-13
ENSG00000117410	ATP6V0B	0.479535962	2.58E-11
ENSG00000147416	ATP6V1B2	0.437789019	5.51E-09
ENSG00000128609	NDUFA5	0.381648881	0.0008
ENSG00000204370	SDHD	0.377797199	0.0003
ENSG00000198712	COX2	0.322901415	0.005
ENSG00000073578	SDHA	-0.159341647	0.04
ENSG00000023228	NDUFS1	-0.210017339	0.002
ENSG00000112695	COX7A2	-0.211631763	0.02
ENSG00000156467	UQCRB	-0.2162753	0.03
ENSG00000167792	NDUFV1	-0.222597617	0.007
ENSG00000135390	ATP5MC2	-0.222831627	0.002
ENSG00000143252	SDHC	-0.24144827	0.02
ENSG00000160194	NDUFV3	-0.242343108	0.004
ENSG00000131100	ATP6V1E1	-0.274149444	2.61E-05
ENSG00000189043	NDUFA4	-0.277227615	0.0001
ENSG00000176340	COX8A	-0.278696825	0.001
ENSG00000165264	NDUFB6	-0.28300413	0.005

ENSG00000213619	NDUFS3	-0.29075047	0.001
ENSG00000178741	COX5A	-0.297956503	9.27E-05
ENSG00000110719	TCIRG1	-0.300385483	0.002
ENSG00000090266	NDUFB2	-0.301527793	0.001
ENSG00000174886	NDUFA11	-0.318594422	0.001
ENSG00000169021	UQCRFS1	-0.319382279	2.33E-05
ENSG00000130414	NDUFA10	-0.320916351	3.89E-05
ENSG00000131143	COX4I1	-0.336149045	2.75E-07
ENSG00000158864	NDUFS2	-0.343567449	4.23E-06
ENSG00000139180	NDUFA9	-0.357605693	0.0006
ENSG00000004779	NDUFAB1	-0.360587197	2.52E-05
ENSG00000184076	UQCR10	-0.375287362	3.00E-06
ENSG00000116459	ATP5PB	-0.379803823	9.50E-08
ENSG00000126267	COX6B1	-0.388197692	3.84E-08
ENSG00000100554	ATP6V1D	-0.405728961	7.75E-08
ENSG00000131495	NDUFA2	-0.419069805	1.65E-06
ENSG00000115286	NDUFS7	-0.422028053	0.0002
ENSG00000136888	ATP6V1G1	-0.428892928	2.02E-09
ENSG00000183648	NDUFB1	-0.432392964	0.0003
ENSG00000154518	ATP5MC3	-0.445435152	1.00E-08
ENSG00000125356	NDUFA1	-0.459056555	2.84E-07
ENSG00000179091	CYC1	-0.468189203	1.54E-09
ENSG00000127540	UQCR11	-0.475454692	4.04E-09
ENSG00000110955	ATP5F1B	-0.49393262	1.83E-16
ENSG00000099795	NDUFB7	-0.509075592	1.10E-09
ENSG00000168653	NDUFS5	-0.520875625	8.23E-11
ENSG00000159720	ATP6V0D1	-0.540902137	6.67E-12
ENSG00000147123	NDUFB11	-0.542107011	1.51E-08
ENSG00000178127	NDUFV2	-0.542969651	0.006
ENSG00000124172	ATP5F1E	-0.574462256	1.96E-13
ENSG00000152234	ATP5F1A	-0.606022757	1.10E-17
ENSG00000140990	NDUFB10	-0.613088113	6.15E-12
ENSG00000119013	NDUFB3	-0.65837698	2.66E-11

ENSG00000169020	ATP5ME	-0.663226048	5.18E-18
ENSG00000159199	ATP5MC1	-0.712677613	1.05E-19
ENSG00000241468	ATP5MF	-0.735984185	1.21E-15
ENSG00000099624	ATP5F1D	-0.737422752	3.56E-17
ENSG00000169429	CXCL8	-0.820462236	3.09E-17
ENSG00000164405	UQCRQ	-0.860634282	1.66E-38
ENSG00000198695	ND6	-0.979723775	1.42E-08
ENSG00000143882	ATP6V1C2	-1.543440061	0.005

8.4 Appendix 4: GO analysis of 3' RNA-seq data

Pathway	*P value*	*Genes*
GO:0042273~ribosomal large subunit biogenesis	1.39E-06	RPL35A, RPL14, RPL7, NPM1, RPL26, RPL5, RPL11, RPL24
GO:0042274~ribosomal small subunit biogenesis	6.22E-05	RPS19, RPS16, RPS17, RPS14, NPM1, RPS6, RPS24
GO:0006414~translational elongation	3.59E-40	RPL18, RPL36A, RPL19, RPL14, RPL13, RPS2, RPS3, RPLP0, RPL10, FAU, RPL11, RPL12, RPS27A, RPL35A, RPS18, RPS19, RPS16, RPS17, RPS14, RPS12, RPS13, RPS11, EEF1D, EEF1B2, RPL27A, RPL35, RPL37, RPL38, RPL30, RPS27, RPL32, RPS29, RPL7, RPL6, RPL9, RPL34, RPL8, RPL3, RPL5, RPL10A, RPL7A, RPL4, RPS20, RPS23, RPS24, EEF1A1, EEF1A2, RPL26, RPL27, RPL23A, RPL24, RPS6, RPS5, RPS8, RPL23, RPL18A, RPL13A, RPL22, RPL21, RPL37A
GO:0006749~glutathione metabolic process	2.23E-05	HAGH, GGT5, GPX1, CTH, GCLC, ALDH5A1, GPX4, IDH1, SOD1, MGST1, SOD2
GO:0006958~complement activation, classical pathway	3.18E-05	C8A, C8B, MBL2, CD55, C3, CLU, C5, C1R, C1S, CFI, C8G
GO:0034599~cellular response to oxidative stress	1.86E-06	EPAS1, DUOX2, DUOX1, PRDX5, SOD1, TPM1, SOD2, APOA4, PYCR1, GPX1, LONP1, HIF1A, HMOX1, GPX4, ETV5
GO:0006956~complement activation	8.15E-06	MBL2, MASP1, C3, C5, CLU, C1R, C1S, C8G, CFHR1, C8A, C8B, CD55, CFH, CFI
GO:0055072~iron ion homeostasis	4.64E-05	TTC7A, HAMP, HPX, HMOX1, TFR2, NDFIP1, HP, CP, SOD1, FTH1, SOD2, FTL
GO:0002526~acute inflammatory response	4.48E-10	MBL2, YWHAZ, MASP1, C3, CLU, C5, C1R, C1S, AHSG, CFHR1, APOA2, ITIH4, CFH, CFI, FN1, F12,

		EPHX2, SAA4, IGF2, IL6R, SIGIRR, C8G, C8A, ORM1, C8B, CD55, F3, F2, ORM2
GO:0009896~positive regulation of catabolic process	5.09E-05	APOA4, PPARA, APOA2, HIF1A, GCLC, APOA5, NDFIP1, MDM2, GJA1, IGF2, PNPLA2, ANGPTL3, FURIN, BARD1
GO:0007015~actin filament organization	1.09E-06	ARHGEF2, LIMA1, ACTC1, ACTN4, CALD1, FSCN1, NEDD9, ACTN1, TTN, PROX1, FLNA, PPP1R9A, RND1, PAK2, CFL1, RHOA, RHOF, DBN1, PLS3, LCP1
GO:0002449~lymphocyte mediated immunity	3.13E-06	MSH6, MBL2, ICAM1, C3, CEBPG, CLU, C5, C1R, C1S, C8G, C8A, C8B, CD55, TUBB, BCL3, FAS, HSPD1, CFI, RAB27A
GO:0019724~B cell mediated immunity	5.42E-05	MBL2, MSH6, C3, CLU, C5, C1R, C1S, C8G, C8A, C8B, CD55, BCL3, HSPD1, CFI, FAS
GO:0006412~translation	1.99E-21	EIF6, RPL18, RPL36A, RPL19, RPL14, RPL13, RBM3, EIF2A, RPS27L, RPL22L1, RPS2, RPS3, EIF4EBP2, RPLP0, RPL10, FAU, RPL11, RPL12, POLG2, EIF2B2, RPS27A, RPL35A, RSL1D1, RPS18, RPS19, RPS16, DHX29, RPS17, RPS14, RPS12, RPS13, RPS11, EEF1D, NACA, EEF1B2, RPL27A, RPL35, RPL37, IGF2BP3, RPL38, DTD1, EIF3D, RPS27, RPL30, RPL7, RPL32, RPS29, EIF3H, MRPL14, RPL6, RPL34, RPL9, RPL8, RPL3, RPL5, RPL4, RPL10A, RPL7A, RSL24D1, RPS20, RPS23, RPS24, EEF1A1, EEF1A2, RPL26, RPL27, RPL24, RPL23A, ETF1, RPS6, RPS5, RPS8, EIF4B, MRPL24, RPL23, RPL18A, GSPT1, RPL13A, RPL22, RPL21, RPL37A
GO:0002443~leukocyte mediated immunity	4.86E-06	MSH6, MBL2, ICAM1, YWHAZ, C3, CEBPG, CLU, C5, C1R, IL6R, C1S, C8G, C8A, C8B, CD55, TUBB, BCL3, FAS, HSPD1, CFI, RAB27A

8.5 Appendix 5: Promoter regions of genes with both altered mRNA expression and enrichment of 5hmC

Ensembl ID	Gene symbol	mRNA expression fold change	5hmC fold change
ENSG00000181418	DDN	3.57	-2.84
ENSG00000105131	EPHX3	2.38	-3.82
ENSG00000086619	ERO1B	1.65	4.05
ENSG00000147408	CSGALNACT1	1.33	3.09
ENSG00000149927	DOC2A	1.26	-2.49
ENSG00000142156	COL6A1	1.24	3.13
ENSG00000137752	CASP1	0.98	2.44
ENSG00000119900	OGFRL1	0.71	2.21
ENSG00000145247	OCIAD2	0.71	3.02
ENSG00000183044	ABAT	-1.26	3.04
ENSG00000144395	CCDC150	-1.50	3.36
ENSG00000167874	TMEM88	-2.02	3.60